

Nelson Colossi

MODELOS PARAMÉTRICOS DE CUSTOS PARA PROJETOS DE
SISTEMAS DE ESGOTO SANITÁRIO

Dissertação apresentada ao Programa de Pós-Graduação em
Engenharia de Produção da Universidade Federal de Santa Catarina,
como requisito parcial, para obtenção do grau de Mestre em Engenharia
de Produção

Orientador: Prof. Antônio Edésio Jungles, Dr.

Florianópolis

2002

Nelson Colossi

**MODELOS PARAMÉTRICOS DE CUSTOS PARA PROJETOS DE
SISTEMAS DE ESGOTO SANITÁRIO**

Esta dissertação foi julgada e aprovada para a obtenção do grau de
Mestre em Engenharia de Produção no **Programa de Pós-
Graduação em Engenharia de Produção** da Universidade Federal de
Santa Catarina

Florianópolis, 18 de setembro de 2002.

Prof. Edson Pacheco Paladini, Dr
Coordenador do Programa

BANCA EXAMINADORA

Prof. Antônio Edésio Jungles, Dr
Universidade Federal de Santa Catarina
Orientador

Prof. Emilio Araújo Menezes, Dr
Universidade Federal de Santa Catarina

Prof. Maurício Luiz Sens, Dr
Universidade Federal de Santa Catarina

A Ugeri Colossi (in memorian) e Ondina Ronconi Colossi, meus pais, por tudo aquilo que fizeram ao longo de minha trajetória.

A minha irmã, Terezinha, pelo apoio constante.

A meus filhos, Fabio, Tatiana e Guilherme, pela compreensão diante da pouca dedicação que lhes tenho oferecido.

Agradecimentos

Ao professor Antônio Edésio Jungles, orientador deste trabalho, por sua amizade, por suas opiniões e pelo incentivo.

À Universidade Federal de Santa Catarina, ao Departamento de Engenharia de Produção e aos seus professores que contribuíram para aumentar o meu conhecimento e possibilitar o desenvolvimento desta pesquisa.

À Companhia Catarinense de Águas e Saneamento (CASAN) pela sua preocupação com aperfeiçoamento dos funcionários.

Aos colegas e profissionais da CASAN que facilitaram meu acesso às informações e aos dados importantes a este trabalho, e àqueles que me acompanharam ao longo da preparação da dissertação e nas oportunidades de troca de informações e experiências.

A todos os que direta ou indiretamente contribuíram para a realização desta pesquisa

“No que as leis da Matemática se referem à realidade, elas não são precisas e, quando são precisas, elas não se referem à realidade”.

Albert Einstein

Resumo

COLOSSI, Nelson. **Modelos Paramétricos de Custos para Projetos de Sistemas de Esgoto Sanitário**. 2002. 137f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis.

Estimativas de custos adequadas e precisas são essenciais na avaliação, análise e tomada de decisão para investimentos no setor de saneamento, especificamente em sistemas de esgoto sanitário. Uma investigação, em Santa Catarina, tem demonstrado a falta de um referencial de custo satisfatório para planejamento, programação e estudos de concepção de projetos de esgoto. Portanto, esta pesquisa concentra-se na análise de uma metodologia alternativa como ferramenta adequada, denominada estimativa paramétrica de custo. Investigações de dados históricos foram realizadas em cerca de cinquenta projetos de sistema de esgoto sanitário existentes. Nesses projetos, os dados foram coletados de componentes básicos que configuram cada sistema, tais como, rede coletora, ligações prediais, estação elevatória, emissário e estação de tratamento de esgoto. Os dados históricos coletados constituem os custos e as características técnicas, que possam se relacionar positivamente a esses custos, de tal modo, que expliquem as relações paramétricas de custo de cada componente de sistema. Análises de regressão múltipla linear foram utilizadas para o ajuste de modelos estatísticos das relações paramétricas de custo. O modelo de custo da rede coletora ajustado tem a variável custo explicada pelas variáveis independentes, área de abrangência do projeto da rede e densidade populacional da área. Já o modelo de custo de ligações prediais apresenta bom ajuste quando explanado pelo número de ligações. A estação elevatória tem seu modelo explicado pelos direcionadores de custo: vazão de projeto e altura manométrica. O ajuste do modelo de emissário é explicado pela vazão de projeto e pela sua extensão. Entretanto, no caso da estação de tratamento, devido à pequena quantidade de dados existentes, o modelo tem ajuste pobre e não apresenta boa significância estatística. Os bons valores, obtidos para os Coeficientes de Determinação (R^2) de cada modelo construído, superiores a 93%, mostram a alta percentagem de variação nos custos explicada pelas variáveis

independentes. A validação desses modelos paramétricos de custo vem confirmar a adequada precisão deles na estimativa de custo de projetos. Desse modo, os modelos justificam os objetivos gerais e específicos propostos, tanto na qualidade dos resultados quanto na facilidade de obtenção de estimativas de custo, já nas fases iniciais de um projeto de sistema de esgoto sanitário.

Palavras-chave: modelo paramétrico de custo, estimativa de custo, sistema de esgoto sanitário, relações de estimativa de custo, análise de regressão

Abstract

COLOSSI, Nelson. **Modelos Paramétricos de Custos para Projetos de Sistemas de Esgoto Sanitário**. 2002. 137f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis.

Appropriate and accurate costs estimations are essential in the evaluation, analysis and decision taking for investments into sanitation field, specifically in sanitary sewer systems. An investigation in the state of Santa Catarina has demonstrated the lack of a satisfactory cost reference for the planning, programming and studies of sewage projects conception. Therefore, this research is centered on the analysis of an alternative methodology as an appropriate tool, denominated parametric cost estimating. Investigations of historical data are made in about fifty wastewater collection and treatment existent projects. On these, data is collected about the basic components that configure each system, such as gravity collector sewer, building connections, pumping station, force main and wastewater treatment plant. The collected historical data constitutes the costs and the technical characteristics that may be related positively with those costs, in such way as to explain the parametric cost estimating relationship of each system component. Linear multiple regression analysis are used for the fitting of statistical models for parametric cost estimating relationships. The adjusted cost model of the gravity collector sewer has its cost variable explained by the independent variables, also denominated cost drivers, such as the area covered by the sewer project and density population of the area. The cost model of building connections is better fitted when explained by the number of connections. The pumping station, on the other hand, has its model explained by the cost drivers: design flow and pumping head. The fit model of force main is explained by the design flow and for its extension. However, for wastewater treatment plant, due to small amount of existent data, the model presents poor fitting and does not good statistical significance. The good values acquired by the Coefficients of Determination (R^2) for each built model, superior to ninety three percent, demonstrate the high percentage of variation in the costs explained by the independent variables. The validation of these parametric cost models corroborates

their appropriate precision for projects cost estimation. Finally, the models justify the proposed general and specific objectives, not only in the quality of the results but also in the easiness of obtaining cost estimates as early as in the initial phases of a sanitary sewer system project.

Key-words: parametric cost model, cost estimates, sewage systems, cost estimating relationships, regression analysis

Sumário

LISTA DE FIGURAS.....	I
LISTA DE TABELAS	II
1 INTRODUÇÃO	14
1.1 Justificativa	16
1.2 Objetivos.....	16
1.2.1 Objetivo Geral	16
1.2.2 Objetivos Específicos.....	17
1.3 Limitações do Trabalho.....	17
1.4 Estrutura do Trabalho	18
2 REVISÃO DE LITERATURA.	20
2.1 Custos	20
2.1.1 Categorias de Custos	20
2.2 Princípios e Métodos de Custeio	21
2.2.1 Métodos de Custeio	22
2.3 Estimativas de Custo.....	23
2.4 Estimativas Paramétricas de Custo.....	24
2.5 Etapas para o Desenvolvimento do Modelo Paramétrico de Custo.....	29
2.6 Identificação e Coleta, e Avaliação de Dados.....	29
2.7 Normalização de Dados	31
2.8 Análise de Direcionadores de Custo.....	33
2.9 Processo de Desenvolvimento de Relações Paramétricas	34
2.10 Modelação Matemática da Relação Paramétrica de Custo.....	36
2.10.1 Fator e Razão	36
2.10.2 Método Gráfico	36
2.10.3 Análise de Regressão.....	37
2.11 Teste de significância da Relação Paramétrica de Custo	43
2.12 Validação do Modelo Paramétrico.	46
2.13 Modelos Paramétricos de Custos Aplicados na Construção e no Saneamento.....	47
3 METODOLOGIA	57
3.1 Caracterização da Pesquisa	57
3.1.1 Pesquisa Aplicada.....	57
3.1.2 Pesquisa Quantitativa	57
3.1.3 Pesquisa Documental e Experimental.....	57
3.2 Fluxograma do Processo de Desenvolvimento de Modelos Paramétricos de Custo.....	58
3.2.1 (A) – Oportunidade de Identificar Modelos.	59
3.2.2 (B) - Coleta de Dados de Projetos de Sistemas de Esgoto Sanitário.....	59
3.2.3 (C) - Normalização dos Dados.....	62
3.2.4 (D) - Reconhecimento e Análise de Direcionadores de Custo.	63
3.2.5 (E) - Desenvolvimento das Relações do Modelo Paramétrico de Custo.....	67
3.2.6 (F) – Modelação Matemática.	71
3.2.7 (G) –Testes de Significância do Modelo Ajustado.	76
3.2.8 (H) – Validação do Modelo.....	77
4 RESULTADOS.....	79
4.1 Modelo de Custo de Rede Coletora de Esgoto Sanitário	79
4.2 Modelo de Custo de Ligação Predial de Esgoto Sanitário.	90
4.3 Modelo de Custo de Estação Elevatória de Esgoto Sanitário.....	97
4.4 Modelo de Custo de Emissário de Esgoto Sanitário.....	105
4.5 Modelo de Custo de Estação de Tratamento de Esgoto Sanitário	115

4.6	Análise dos Resultados dos Modelos	118
4.7	Validação dos Modelos	122
4.7.1	Comparativo para um Sistema entre o Custo de Projeto e o Estimado pelos Modelos	125
5	CONCLUSÕES E RECOMENDAÇÕES	127
5.1	Conclusões do Trabalho	127
5.1.1	Conclusões Gerais.....	127
5.1.2	Conclusões Específicas.....	128
5.2	Como Utilizar os Modelos	129
5.3	Recomendações para Pesquisas Futuras	130
6	REFERÊNCIAS	132

Lista de Figuras

Figura 3.1. Fluxograma de desenvolvimento de modelo paramétrico de custo.	58
Figura 4.1. Custos de construção de redes coletoras em função de área de abrangência (ha) e densidade populacional (hab/ha).	81
Figura 4.2. Custos de rede coletora em função de área de abrangência.	81
Figura 4.3. Custos de redes coletoras em função de densidade populacional (hab/ha).	82
Figura 4.4. Resíduos dos custos estimados de redes coletoras.	86
Figura 4.5. Resíduos dos custos estimados de rede coletora em relação à variável área de abrangência.	87
Figura 4.6. Resíduos dos custos estimados de rede coletora em relação à variável densidade populacional.	87
Figura 4.7. Gráfico da probabilidade normal dos resíduos dos custos estimados de redes coletoras.	88
Figura 4.8. Custos estimados e observados de redes coletoras.	89
Figura 4.9. Custos estimados e observados de redes coletoras através de índice (número de ordem dos dados).	89
Figura 4.10. Custos de ligações prediais em função de número de ligações.	91
Figura 4.11. Resíduos dos custos estimados de ligações prediais.	94
Figura 4.12. Resíduos dos custos estimados relacionados à variável n° de ligações.	94
Figura 4.13. Resíduos em função de índice (n° de ordem dos dados).	95
Figura 4.14. Gráfico de probabilidade normal dos resíduos dos custos estimados de ligações prediais.	95
Figura 4.15. Custos estimados e observados de ligações prediais.	96
Figura 4.16. Custos estimados e observados de ligações prediais através de índice (n° de ordem dos dados).	96
Figura 4.17. Custos de estações elevatórias em função de capacidade e altura manométrica.	99
Figura 4.18. Custos de estações elevatórias em função de capacidade.	99
Figura 4.19. Custos de estações elevatórias em função de altura manométrica.	100
Figura 4.20. Custos estimados e respectivos custos observados para estações elevatórias.	103
Figura 4.21. Resíduos dos custos estimados de estações elevatórias relacionados com a variável capacidade.	103
Figura 4.22. Resíduos dos custos estimados de estações elevatórias em relação à variável altura manométrica.	104
Figura 4.23. Resíduos dos custos estimados de estações elevatórias.	104
Figura 4.24. Gráfico da probabilidade normal dos resíduos dos custos estimados das estações elevatórias.	105
Figura 4.25. Custos estimados e observados de estações elevatórias através de índice (número de ordem dos dados).	105
Figura 4.26. Custos de emissários em função de vazão e extensão.	108
Figura 4.27. Custos de emissários em função de vazão de projeto.	108
Figura 4.28. Custos de emissários em função de sua extensão.	109
Figura 4.29. Resíduos dos custos estimados de emissários.	112
Figura 4.30. Resíduos dos custos estimados de emissários em relação à vazão de projeto.	112
Figura 4.31. Resíduos dos custos estimados de emissários em relação à extensão.	113
Figura 4.32. Gráfico de probabilidade normal dos resíduos dos custos estimados de emissários.	113
Figura 4.33. Custos estimados em função de custos observados de emissários.	114
Figura 4.34. Custos estimados e observados de emissários relacionados aos índices (numeração de ordem dos dados).	114
Figura 4.35. Influência dos fatores de economia de escala no modelo de custo de rede coletora.	119
Figura 4.36. Influência do fator de economia de escala no modelo de custo de ligações prediais.	119
Figura 4.37. Influência dos fatores de economia de escala no modelo de custo de estação elevatória.	120
Figura 4.38. Influência dos fatores de economia de escala no modelo de custo de emissário.	120

Lista de Tabelas

Tabela 2.1. Equações Paramétricas de custo para estação elevatória.....	50
Tabela 3.1. Discriminação de custos de construção por componentes de sistema de esgoto sanitário.	61
Tabela 3.2. Análise comparativa entre os possíveis ajustes do custo de rede coletora e seus direcionadores.....	68
Tabela 3.3. Análise comparativa entre os possíveis ajustes do custo de ligações prediais e seus direcionadores.....	69
Tabela 3.4. Análise comparativa entre os possíveis ajustes do custo de estação elevatória e seus direcionadores.....	69
Tabela 3.5. Análise comparativa entre os possíveis ajustes do custo de emissário e seus direcionadores.....	70
Tabela 3.6. Análise comparativa entre os possíveis ajustes do custo de estação de tratamento e seus direcionadores.....	70
Tabela 4.1. Dados de projetos de redes coletoras.....	80
Tabela 4.2. Ajuste do modelo de custo de rede coletora.....	84
Tabela 4.3. Análise da variância do ajuste do modelo.....	85
Tabela 4.4. Dados de projetos de ligações prediais.....	91
Tabela 4.5. Ajuste do modelo de custo de ligações prediais.....	92
Tabela 4.6. Análise de Variância de Ajuste do Modelo.....	93
Tabela 4.7. Análise da variância das variáveis independentes.....	93
Tabela 4.8. Dados de projetos de estações elevatórias de esgoto sanitário.....	98
Tabela 4.9. Ajuste do modelo de custo de estações elevatórias.....	101
Tabela 4.10. Análise da Variância do Ajuste do Modelo.....	102
Tabela 4.11. Dados de projetos de emissários.....	107
Tabela 4.12. Ajuste do modelo de custo de emissários.....	110
Tabela 4.13. Análise da variância do ajuste do modelo.....	110
Tabela 4.14. Análise da variância das variáveis independentes.....	111
Tabela 4.15. Dados de projetos de estação de tratamento.....	116
Tabela 4.16. Ajuste do modelo de custo de estações de tratamentos.....	117
Tabela 4.17. Análise da variância do ajuste do modelo.....	117
Tabela 4.18. Modelos matemáticos para as relações paramétricas de custo dos componentes de sistema de esgoto sanitário.....	118
Tabela 4.19. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de rede coletora.....	123
Tabela 4.20. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de ligações prediais.....	124
Tabela 4.21. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de estação elevatória.....	124
Tabela 4.22. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de emissário.....	125
Tabela 4.23. Dados de custo de projeto do S.E.S. de Joinville – bacia 3.2.....	125
Tabela 4.24. Custos de projeto e estimados do S.E.S. de Joinville – bacia 3.2.....	126
Tabela 4.25. Custo de projeto do S.E.S. de Joinville – bacia 3.2. e custo estimado atualizado.....	126

1 INTRODUÇÃO

Pesquisadores, planejadores, projetistas e organismos financiadores do saneamento no Brasil, principalmente em Santa Catarina, têm se defrontado com um grande problema: a falta de estudos de custos que possibilitem um referencial para tomada de decisão na avaliação e análise de investimentos no setor de construção do saneamento. Isso ocorre especificamente em Sistemas de Esgoto Sanitário. Frequentemente, busca-se este referencial em simples relações, sem muita precisão, como algumas medidas de projeto que possam explicar os custos, que na maioria das vezes, se apoiam na experiência do profissional que atua no setor (JUNGLES, 1994).

A utilização de índices ou relações de custo, desenvolvidos para outras atividades, na avaliação ou estimativa de custo das obras em projetos de saneamento, tem apresentado, em geral, problemas das mais variadas ordens, como por exemplo:

- incerteza do público alvo quanto ao resultado da avaliação do custo;
- composição inadequada para obras de saneamento;
- parâmetros de referência inadequados;
- unidade básica referencial imprópria;
- proporção inadequada dos insumos da composição do índice;
- distorção na estimativa de custo de projetos;
- e outros não menos importantes.

“Uma estimativa é uma previsão, uma aproximação, que produz informações para decisões empresariais e se apresenta como substituto para a medição real, quando esta não é viável física e economicamente. É considerada precisa, quando é suficientemente próxima ao valor real, de modo que as decisões tomadas com base nessa estimativa são similares às aquelas fundamentadas num ambiente real, caso sua caracterização fosse possível” (CARR *apud* OTERO, 2000).

A determinação adequada dos custos em um projeto é indispensável, bem como sua precisão.

Várias técnicas de estimativa de custo são disponíveis, mas a técnica adequada depende da fase de definição do projeto. A sua escolha depende também da finalidade de uso (PANZETER, 1993).

Sabe-se que o método de estimativa de custo tem origem nas características físicas e tecnológicas do projeto e em determinadas regras e critérios de análise. Assim, ele deve então montar um retrato, com um erro não significativo se comparado àqueles valores de custos que seriam detalhados no projeto e à precisão requerida para o fim ao qual se destina tal estimativa.

Em diversas funções, há a possibilidade de uso de métodos de estimativa os quais mantêm níveis de precisão inferiores ao do custo detalhado, porém compatíveis com grau de exatidão exigido nesses casos. Tais técnicas são caracterizadas pela alta velocidade de processamento de resultados, sendo por isso denominadas expeditas.

A técnica do uso de relações paramétricas proposta neste trabalho, também é considerada expedita, pois utiliza um número reduzido de informações com características relacionadas ao custo do projeto. Essa técnica tem seu processamento baseado em uma única equação para cada projeto.

A estimativa paramétrica de custos de cada parte de sistema de esgoto sanitário, em estudo, encaminha-se justamente nesse sentido, fazendo uso de informações menos detalhadas e normalmente disponibilizadas nas primeiras etapas do estudo de concepção do sistema. Inicialmente, a área de abrangência do projeto, a população, a densidade demográfica, a vazão de projeto e as extensões de ruas, são por exemplo, dados relacionáveis com os custos dos citados componentes.

Por conseguinte, a pesquisa está direcionada de maneira que ao final desta, sejam criados e validados modelos paramétricos de custo de partes componentes de sistema de esgoto sanitário.

O desenvolvimento desta pesquisa trará uma nova abordagem para avaliação de custo de obras de saneamento, desde a concepção da idéia inicial, a organização e o gerenciamento das diversas etapas, a forma da interação interdisciplinar, o levantamento e a obtenção dos diversos insumos que irão compor a aplicação, até os procedimentos de utilização para avaliação do custo das obras de um projeto.

O trabalho terá como base a coleta e a compilação de dados existentes para cada parte de sistema, a avaliação desses dados, o desenvolvimento de relações paramétricas de custo, a modelação matemática e os testes de significância . Por fim, a validação dos modelos de custo o complementar .

1.1 Justificativa

A aplicabilidade do estudo está fundamentada na inexistência de uma metodologia de composição de modelos paramétricos de custo para projetos de esgoto sanitário em Santa Catarina. Fato esse que tem contribuído, em grande parte, para que as avaliações de custo de projetos, destinadas a essa parte do saneamento, apresentem variados problemas de distorções nas estimativas. Como consequência, tem afetado cronogramas financeiros, pedidos de financiamentos, e gerado escassez de recursos empenhados nas obras e principalmente causado incertezas quanto ao custo final do projeto durante a fase de estudos econômicos de alternativas ou de elegibilidade. .

A realidade encontrada é de vários estudos sobre os custos de uma ou de outra parte das obras de um sistema de saneamento, porém faltando um conjunto de procedimentos condizentes entre si e com base científica.

As estimativas de custo disponíveis geralmente originam-se de esforços individuais, ou de pequenos grupos, para atender uma determinada idéia ou necessidade, que geralmente não atinge às expectativas dos usuários na avaliação dos custos de projetos.

A aplicação da metodologia desenvolvida proporcionará um instrumento de apoio na elaboração de projetos de esgoto sanitário, nas suas diversas etapas.

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo geral deste trabalho de pesquisa é desenvolver modelos e relações paramétricas de custo que forneçam resultados confiáveis e do mais alto nível, com menos detalhes de projeto, a um baixo custo e num menor tempo de ciclo que outros métodos tradicionais.

1.2.2 Objetivos Específicos

Os objetivos específicos da pesquisa são:

- Analisar o nível de precisão alcançado por modelos paramétricos de custo nas estimativas de custo de construção de projetos de sistema de esgotos sanitários convencionais de cidades, bairros, loteamentos ou áreas urbanas e de partes componentes desse sistema.
- Possibilitar a elaboração de custos de construção de sistemas de esgoto sanitário e de seus componentes, antes da realização do projeto, por planejadores e por agentes de financiamento.
- Verificar a possibilidade de modelar os custos, em função dos pesos de algumas variáveis, de maneira que o modelo possa explicar adequadamente os custos de construção do sistema ou de cada um de seus componentes.
- Comparar os custos para construção de projetos da Companhia Catarinense de Águas e Saneamento (CASAN), com as estimativas de custo de construção fornecidas pelos modelos desenvolvidos na pesquisa.
- Investigar quais são as variáveis, sejam elas geométricas, quantidades de material ou de capacidade, que melhor explanam os custos.
- Determinar, através de análise estatística, a natureza da relação entre as variáveis dependente e independente, quando uma atenção especial deve ser dada para os modelos lineares de primeira ordem.

1.3 Limitações do Trabalho

As informações e dados apresentados no trabalho são restritos num primeiro momento à Companhia Catarinense de Águas e Saneamento e ao Estado de Santa Catarina. A extrapolação de tais dados para uso fora destes limites deve ser precedida por uma verificação de compatibilidade entre as informações presentes no ambiente citado e as do novo universo que se deseja como objeto de estudo.

Os dados utilizados são fundamentados em orçamentos detalhados de projetos de sistemas de esgoto sanitário, cujo o objetivo é a licitação e posterior execução destes projetos. Tem-se por pressuposto que as composições unitárias e preços de

insumos utilizados, e conseqüentemente os totais acumulados em quantidades e custos de material e mão de obra, são compatíveis com a realidade esperada para o processo de implantação daqueles projetos, retratando valores médios esperados.

O mesmo ocorre com quantidades e custos de serviços, sendo pressuposto que não houve falhas no levantamento destes quantitativos, seja na falta ou excesso de partes de obras, ou no procedimento de cálculo de valores. A empresa, que disponibilizou os dados, mantém procedimentos padronizados para o processo de orçamento, tendo sido produzidos a partir destas regras aqueles que formam as amostras deste trabalho.

A verificação de valores espúrios é feita unicamente a partir da análise estatística, não sendo discutidas diferenças com relação às características dos projetos que repercutam pequenas alterações nos custos. Em alguns casos, entretanto, a definição de diferentes padrões de custo e quantidade para os projetos, é explicada por meio de análises dos dados de custos e de informações gerais sobre os métodos, processos tecnológicos e aspectos geográficos relativos aos projetos.

1.4 Estrutura do Trabalho

A Seção 1 deste trabalho, denominada Introdução, faz uma apresentação do problema e aponta as justificativas para solucioná-lo, a partir da conceituação e determinação de modelos paramétricos de custo dentro da atividade de projeto de sistemas de esgoto sanitário. Também expõe os objetivos gerais e específicos desta pesquisa, as limitações estabelecidas e a caracterização da própria estrutura do trabalho.

Na Seção 2, Revisão de Literatura, faz-se a apresentação da teoria sobre custos, estimativas de custos e relações paramétricas de custo, englobando histórico, evolução, conceituação e procedimentos para uso. Também é mostrada uma revisão sobre as etapas de desenvolvimento das relações paramétricas de custo e de modelos aplicados nos setores da construção e do saneamento.

A caracterização da pesquisa e a metodologia são mostrados na Seção 3, Metodologia. É apresentado um fluxograma do processo de desenvolvimento dos modelos seguido da descrição detalhada de cada etapa dele, visando a boa

interpretação da construção dos modelos dos componentes de sistema de esgoto sanitário. As etapas são identificadas pela proposição dos modelos, coleta e, normalização dos dados, definição dos direcionadores de custo, desenvolvimento das relações entre as variáveis dos modelos, modelação matemática, testes de significância e validação dos modelos.

Os resultados da pesquisa são apresentados na Na Seção 4, Resultados. Lá são identificados, avaliados e normalizados os dados históricos, e descritas as composições de custo das partes de sistema de esgoto sanitário. Os modelos são desenvolvidos, ajustados e testados estatisticamente para cada um desses componentes de sistema. É feita a validação deles através de dados de custo independentes daqueles dados que os formaram, mostrando a qualidade da precisão de suas equações matemáticas. No final, é apresentada a estimativa de custo de um sistema de esgoto sanitário selecionado com exemplo.

A Seção 5, Conclusões e Recomendações, apresenta as principais conclusões produzidas neste trabalho, e aborda os direcionadores de custo, a aplicação dos modelos paramétricos de custo e os padrões de comportamento nas relações paramétricas. São colocadas ainda orientações de como usar os modelos e propostas para pesquisas futuras, fundamentadas nas dificuldades encontradas ao longo deste trabalho e na possibilidade de sua continuação.

2 REVISÃO DE LITERATURA.

Apropriar custos de produtos ou de serviços, buscar formas de redução de seu valor e analisar as várias decisões tomadas pelas empresas são questões que há muito tempo são discutidas. Discutem-nas técnicos, pesquisadores e empresários, com o foco voltado para o gerenciamento dos custos, visando a competitividade, e para a qualidade desses produtos ou serviços.

Segundo KLIEMANN e MULLER (1994), os sistemas de gerenciamento de custos constituem-se de conjuntos de informações gerenciais organizadas para o processo de tomada de decisões, planejamento e controle de atividades empresariais.

2.1 Custos

Vários autores que fazem uso ou pesquisam os custos, como MARTINS (2000), conceituam-no como sendo um gasto relativo a um bem ou a um serviço utilizado na produção de outros bens ou serviços. Ele representa o valor da soma dos insumos (mão-de-obra, materiais e equipamentos) necessários à realização de um dado produto ou serviço, constituindo-se no valor pago pelos insumos.

Dependendo do seu comportamento em relação à produção ou ao modo como ocorre numa empresa, o custo pode ser classificado em categorias.

2.1.1 Categorias de Custos

Os custos podem ser classificados de acordo com a quantidade produzida ou então conforme a facilidade de alocação deles aos produtos ou serviços (ANTUNES JUNIOR, 1998).

Relativamente à quantidade produzida, eles são classificados em (KLIEMANN NETO, 1986):

- **custo fixo** – é aquele montante que ocorre na empresa, independentemente do volume de produção ou de serviço, ou seja, não está relacionado às oscilações na atividade de produção. Por exemplo, custo de pessoal administrativo, aluguéis, gastos em comunicação, seguro, entre outros;

- **custo variável** – é o montante diretamente proporcional às quantidades produzidas, ou seja, passa a ser constante por unidade de produto ou serviço. Por exemplo: materiais, mão-de-obra direta, impostos e taxas, entre outros.

Os custos relacionados com a facilidade de alocação aos produtos ou serviços são classificados em duas categorias, conforme FULLER e ARMSTRONG (1999):

- **custos diretos** – são os gastos industriais que podem ser alocados direta e objetivamente aos produtos. Assim, ao se considerar uma linha de produtos, os custos de materiais e de mão-de-obra envolvidos em sua manufatura são ambos custos diretos;
- **custos indiretos** – são aqueles que, ao apresentarem dificuldade para serem atribuídos diretamente aos produtos ou às atividade produtivas, requerem para sua apropriação critérios de distribuição ou rateio. Por exemplo: ferramentas, trabalhos de apoio, instalações auxiliares, administração e manutenção da obra.

2.2 Princípios e Métodos de Custeio

A combinação conveniente de princípios e métodos de custo, segundo KLIEMANN e ANTUNES JÚNIOR (1990), é o processo de custeio que deve ser estabelecido para possibilitar a obtenção das informações necessárias a uma empresa ou a uma organização.

Conceitos relativos a esses aspectos que formam um sistema de custeio são apresentados na seqüência.

De acordo com BORNIA e SANTOS (1997), os princípios de custeio são filosofias a serem seguidas pelos sistemas de custos, de acordo com o objetivo e o período de tempo aplicados.

Esses princípios de custeio abordam quais os custos que podem ou devem ser atribuídos aos produtos ou serviços.

KLIEMANN e MULLER (1994), por outro lado, reconhecem o princípio de custeio como a forma com que os custos fixos e variáveis são determinados.

Dois princípios são identificados por MARTINS (2000), um é o custeio por absorção parcial e o outro o custeio direto ou variável.

Existe também um terceiro, segundo ORNESTEIN *apud* BORNIA (1995), conhecido como o custeio por absorção total ou integral.

O **custeio por absorção parcial** avalia como custo do produto ou serviço apenas a parcela ideal de custos fixos relativos à efetiva utilização da capacidade instalada da empresa. Enquanto a parte ociosa ou ineficiente da empresa, que representam os custos fixos, é lançada como perdas do período.

Já o **custeio por absorção total ou integral** considera que todos os custos fixos são alocados aos produtos independente do nível de atividade da empresa. Como rateia-se a totalidade dos custos fixos sobre os produtos, esse princípio visa basicamente atender as exigências da legislação para efeito de avaliação de estoque (BORNIA, 1995).

Por fim, o **custeio direto ou variável** atribui ao produto ou serviço apenas os custos variáveis. Esse princípio define os custos fixos como despesas do período de produção.

Com o objetivo de repassar aos produtos ou os serviços os diversos itens de custos de uma empresa surgiu a necessidade do desenvolvimento de métodos de custeio. Principalmente os itens relacionados com os custos indiretos aos produtos (ANTUNES JUNIOR, 1998).

2.2.1 Métodos de Custeio

De acordo com BORNIA (1995), há quatro métodos de Custeio. O método do custo padrão, ou método dos centros de custos, o método da unidade de esforços de produção (UEP), e por último o método ABC.

O **método do custo padrão** é basicamente uma metodologia de custeio relativo apenas à matéria prima, à mão-de-obra direta ou aos insumos mais relevantes, não podendo, segundo BORNIA (1995), ser aplicada a todos os custos da empresa. Seu objetivo principal é fornecer suporte para o controle dos custos da empresa.

No **método dos centros de custos**, também conhecido por método das seções homogêneas, a alocação dos custos se dá em duas fases. Na primeira, a empresa divide-se em centros de custos indiretos, onde os custos incorridos em determinado período são alocados por meio de bases de rateio previamente definidas. Obtêm-se

assim os custos totais do período para cada centro de custos. Na segunda fase, os custos são alocados dos centros produtivos para os produtos, por meio de bases de rateio (BORNIA, 1995).

O **método da unidade de esforço de produção** busca custear os produtos a partir da identificação da intensidade com a qual os produtos utilizam a estrutura produtiva da empresa. A sua operacionalização ocorre através do estabelecimento dos postos operativos e pela determinação do valor da unidade de esforço de produção (BORNIA, 1995).

O **método ABC** ou custeio baseado em atividades, segundo BORNIA (1995), é a metodologia que toma os custos das várias atividades da empresa e analisa o comportamento delas, buscando bases que representem as relações entre os produtos e essas atividades. De acordo com COKINS (1998), o custeio baseado em atividades é um método que faz o rastreamento dos custos dos negócios para as atividades realizadas e verifica como essas atividades estão relacionadas à geração de receitas e ao consumo dos recursos.

2.3 Estimativas de Custo

São vários os fatores envolvidos na decisão para investir em um novo projeto. Um dos mais importantes é a necessidade de uma estimativa de custo precisa, porque a sua superestimativa ou subavaliação pode causar desastres financeiros (PANZETER, 1993). De acordo com a Curva de Freiman apresentada por DASCHBACH e APGAR *apud* PANZETER (1993) ou mesmo a Curva de WALTER (1999), pode-se chegar a três conclusões:

- quanto maior a subavaliação, maior o desembolso ou desperdício real;
- quanto maior a superestimativa, maior o desembolso ou desperdício real;
- quanto mais realista a estimativa, mais econômico é o projeto.

Há muitas técnicas de estimativa de custo e entre as mais freqüentemente utilizadas incluem-se métodos como o analógico (analogous), o detalhado (bottoms-up) e a estimativa paramétrica.

Ao escolher uma metodologia, o analista deve sempre lembrar que estimativa de custo é uma previsão de custos futuros baseados em uma inferência lógica de

dados históricos disponíveis. O tipo de método usado dependerá da adaptabilidade da definição do produto ou do serviço, do nível de detalhe requerido, da disponibilidade dos dados e, de limitações de tempo (LONG, 2000).

A responsabilidade básica do estimador de custo de projeto é selecionar a metodologia de estimativa mais confiável e da maneira mais econômica para a organização. E o grande desafio é desenvolver modelos de custo que determinem estimativas razoáveis antes de se ter, a mão, todos os detalhes de projeto do produto ou serviço (BOOTHROYD, 1998).

Em muitas circunstâncias, técnicas de estimativa paramétrica podem fornecer resultados confiáveis e do mais alto nível, com menos detalhes de projeto, a um baixo custo e num menor tempo de ciclo que outros métodos tradicionais (DEPARTMENT OF DEFENSE, 1999).

Por conseguinte, esta pesquisa encontrou, na estimativa paramétrica, a metodologia mais apropriada e nesta seção fará um resumo deste procedimento, desde sua origem, passando pela sua definição, desenvolvimento, até sua aplicabilidade.

2.4 Estimativas Paramétricas de Custo.

LARGE (1981) notou em numerosas publicações da Rand Corporation que a estimativa paramétrica não é uma técnica recente. Observando Quéops e imaginando a construção da grande pirâmide, ele poderia expressar os custos de construção como $Y = aX$, onde X é igual a algum parâmetro como blocos de granito ou número de operários.

“Muitas estimativas, através de séculos, e até hoje, são baseadas em custo por libra, custo por pé, custo por barril ou alguma outra unidade simples de medida - todos representam simples modelos paramétricos de custos”(LARGE,1981).

No entanto, LARGE (1981) acredita que modelos paramétricos não eram reconhecidos ou categorizados como modelos de estimativa de custos até o início de 1960, quando a terminologia tornou-se parte do vocabulário da comunidade especialista do Departamento de Defesa dos Estados Unidos da América.

Evidentemente, os modelos paramétricos mostraram-se robustos e viáveis como novos paradigmas no mundo da estimativa de custo, dentro do Departamento de Defesa (ROBBINS e SMITH, 1999).

Segundo ROBBINS e SMITH (1999), analistas de custo que usam estimativas paramétricas há décadas, tornaram-se defensores delas, não só pelas experiências passadas mas também por causa das aplicações. O uso de modelos paramétricos pode gerar estimativas precisas e confiáveis e mais breves que os métodos mais completos, cuja análise detalhada ocorre em todos os itens de serviço, em todos os componentes e composições.

Para ilustrar, um estudo de COLLINS (1996) mostra que cinco empresas da indústria aeronáutica, dos Estados Unidos da América, constataram que as estimativas paramétricas davam resposta 25 a 40% mais rápido e 20 a 35% mais barato e tão confiáveis ou melhor e mais precisas que outros métodos de estimativa, tais como o “analogous” e “bottoms-up”.

Portanto, determinando mudanças de paradigma, vem a necessidade de entender completamente a estimativa paramétrica de custo. E, para aceitar a responsabilidade de usá-la apropriadamente na avaliação de custos, desenvolve-se o interesse em aprender a funcionalidade e os procedimentos da metodologia.

ROBBINS e SMITH (1999) constataam que o significado da estimativa paramétrica de custo pode diferir extremamente de analista para analista e de aplicação para aplicação. Conseqüentemente, para clarear qualquer ambigüidade, esta seção define, descreve, compara e mostra contrastes de alguns fundamentos básicos relativos a estimativa paramétrica.

Dentre muitas definições desta estimativa, a mais sucinta e mais abrangente vem do Glossário de “Society of Cost Estimating and Analysis (SCEA)”:

“Estimativa Paramétrica de Custo é uma metodologia que usa relações estatísticas entre custos históricos e outras variáveis do programa, como características físicas ou de desempenho do sistema, medidas de produção, mão de obra utilizada, etc. Uma técnica que emprega uma ou mais relações de estimativa de custo para medir os custos associados com o desenvolvimento, fabricação e/ou modificação de um artigo de fim específico, baseado nas suas características físicas, técnicas ou outras (SCEA, 1991)”.

Essa definição é aceita pelos analistas que fazem estimativas paramétricas de custo há décadas. Entretanto, embora sua unanimidade, ela não cobre todas as situações. Há que se considerar, também, as seguintes definições achadas em pesquisa superficial:

POLLOCK (1981), em seu artigo “Parametric Primer”, define paramétrica como “...uma simples relação usada para custos de projeto, ou um sistema complexo de estimativa suportado por computadores e uma base de dados extremamente grande...”

Outra definição vem de ROSE *apud* PANZETER (1993), que entende estimativa paramétrica como: “a interpolação ou extrapolação limitada dos dados de custo onde os custos são apresentados como função de variáveis físicas ou relações produzidas”.

Para MENDEL *apud* PANZETER (1993), a estimativa paramétrica consiste em uma técnica de avaliação conceitual que usa quantitativos históricos, custos unitários e critérios de projetos similares para desenvolver um modelo de estimativa de custo.

O “Volume II Aeronautical” de AFMC COST ESTIMATING HANDBOOK SERIES (1982) define a estimativa paramétrica de custo como “uma técnica que emprega uma ou mais relações de estimativa de custo, envolvendo uma coleção de dados históricos relevantes e, através do uso de técnicas matemáticas, relacionando-os com a área a ser estimada”.

U.S. ARMY CORPS OF ENGINEERS (1998), em seu Manual de Instruções Técnicas, aborda a estimativa paramétrica como uma metodologia computacional que usa fatores baseados em parâmetros de engenharia desenvolvidos de bancos de dados históricos de custos, de práticas de construção, e de tecnologia de construção e de engenharia. Esses fatores incluem propriedades físicas que descrevem características de definição de projeto, como tamanho, tipo de construção, tipo de fundação, materiais externos, tipo e material de telhado, número de andares, definição do espaço funcional, definição da utilidade da dependência, etc.

Por outro lado, ABRANOVIC (1997) define métodos paramétricos como distribuição dependente que tira vantagem da informação da amostra para produzir os resultados mais confiáveis e com a maior precisão que as hipóteses possam

assegurar. Os métodos paramétricos podem ser considerados como 100% eficientes quando as hipóteses são satisfeitas.

SHESKIN (1997) ainda reforçou as definições de seus colegas acadêmicos pela diferenciação entre testes paramétricos e não paramétricos. Sheskin definiu os testes como sendo baseados, primeiramente, no nível de medida representado pelos dados analisados. Como regra geral, testes estatísticos inferenciais, que avaliam dados de relação ou de intervalo, são definidos como testes paramétricos. Com base no entendimento geral, deduz-se que, enquanto não há razão para crer que uma ou mais hipóteses de um teste paramétrico tenha sido violada e o nível de medida para os dados ainda seja intervalar ou relacional, os dados devem ser avaliados com o teste paramétrico apropriado.

Conseqüentemente, as definições de estimativa paramétrica podem ser resumidas como a metodologia que usa características físicas ou de desempenho para avaliar um parâmetro relacionado, a qual pode ser horas ou custos, com hipóteses que consideram o banco de dados históricos o elemento definidor.

Assim, não obstante a definição preferida, as estimativas paramétricas podem também ser identificadas como modelos computacionais e bancos de dados. Portanto, a seleção de modelos e a base de dados são dependentes daquilo que está para ser avaliado. Seja usando um modelo existente seja desenvolvendo um novo, o banco de dados é que define.

Segundo SMITH e ROBBINS (1999), entre os usuários de paramétricas, estão incluídos os analistas de custo do Departamento de Defesa dos Estados Unidos da América, seus fornecedores da indústria aeronáutica e aeroespacial, todos os estatísticos e aquele que faz estimativas paramétricas de custo. Assim como, qualquer um que use equações matemáticas para fazer uma previsão pode, compreensivelmente, ser um usuário de paramétricas.

A metodologia paramétrica é usada em situações específicas. Ou seja, os analistas as usam na fase inicial de aquisição de um programa ou projeto, porque ela é, com freqüência, a única técnica de fazer estimativas quando não se dispõe de informações detalhadas. Utiliza-se também na fase final para checar outros métodos de estimativa utilizados, como por exemplo, a estimativa completa e

detalhada (grass roots) (MENÉNDEZ e MORENO, 1995) (DEPARTMENT OF DEFENSE, 1999).

Em qualquer lugar, onde se faça estimativas de custo, as paramétricas são usadas, principalmente em instituições governamentais, consultorias especializadas em custos, indústrias em geral, na seleção da fonte de serviço ou produto, na solicitação de análise e desenvolvimento de propostas ou mesmo na preparação e desenvolvimento de modelos. Até onde se queira um tempo de resposta mais curto e a necessidade de precisão seja importante.

A situação e a função específica são a razão para se usar estimativas paramétricas. Outra razão é que elas geram estimativas de custo mais rápidas e mais baratas e ainda tão ou mais confiáveis que outras metodologias mais detalhadas (COLLINS, 1996) (MENÉNDEZ e MORENO, 1995).

Existem várias técnicas paramétricas que os seus especialistas usam para calcular custos. Relações paramétricas de custo (RPC) e os modelos paramétricos estão entre essas técnicas e são as metodologias mais utilizadas.

A base da metodologia de estimativa paramétrica são as relações paramétricas de custo. Dentro das RPC, estabelece-se uma relevante ligação entre o custo e uma determinada característica técnica do produto ou serviço. Essa é expressa por uma equação matemática. Assim, as relações paramétricas de custo podem significar especificamente expressões matemáticas de graus variados de complexidade que estabelecem o custo como função de um ou mais parâmetros técnicos, como as características físicas ou operacionais. Essas características, que efetivamente definem o custo a ser estimado, denominam-se direcionadores de custo (LONG, 2000). As RPC podem relacionar custo com custo, e custo com variável característica. Um exemplo de relação de custo função de custo é usar horas de produção (variável independente) para estimar os custos de material refugado (variável dependente). A noção básica é que o custo de um elemento é usado para estimar o custo de outro elemento. Quando as RPC são descritas como uma relação custo função de características significa, por exemplo, estimar o custo de produção usando o peso de um item como característica (LONG, 2000).

Geralmente, o detalhamento do custo total em parcelas menores, com a utilização de direcionadores de custo para cada uma de suas partes, possibilita uma

redução do erro intrínseco ao método de estimativa. Assim, pode-se fazer uma composição de relações paramétricas de custo, que, quando utilizadas, conjuntamente estabelecem uma estimativa completa de custo. Este conjunto de relações é denominado um modelo paramétrico.

Os modelos paramétricos são mais complexos porque eles incorporam muitas equações, hipóteses, lógica, e variáveis que descrevem e definem a situação particular sendo estudada e estimada (DEPARTMENT OF DEFENSE, 1999).

2.5 Etapas para o Desenvolvimento do Modelo Paramétrico de Custo.

Várias são as etapas que devem ser consideradas para se obter uma estimativa paramétrica de custo que forneça resultados com confiabilidade e precisão. Portanto, quanto melhor ajustado o modelo paramétrico estabelecido, mais próximo o custo estimado estará do custo real ao final do projeto.

A série de etapas necessárias para o modelo deve corresponder aos seguintes passos:

- identificação, coleta e avaliação de dados históricos relevantes;
- normalização dos dados;
- reconhecimento e análise de direcionadores de custo;
- processo de desenvolvimento de relações paramétricas de custo;
- modelação matemática da relação paramétrica de custo;
- teste de significância da relação paramétrica de custo; e
- validação do modelo paramétrico.

2.6 Identificação e Coleta, e Avaliação de Dados.

O primeiro passo é analisar o conjunto de dados disponível. Isso envolve identificar e coletar dados históricos relevantes, sejam de arquivos de contratos, de informações de custos e cronogramas ou de origens técnicas diversas.

As atividades de coleta e sistematização de dados, bem como a avaliação deles com relação a sua funcionalidade, são passos críticos dentro do processo de estimativa, consumindo tempo e afetando diretamente a precisão dos resultados finais. A correção e confiabilidade da base de dados utilizada são fatores determinantes para a qualidade da estimativa desejada (DEPARTMENT OF DEFENSE, 1999).

Em condições ideais, os dados necessários à estimativa paramétrica desejada devem ser extraídos de informações de projetos completos. Tais informações incluem as de custo, as de informações técnicas e as de cronograma do projeto (ROBBINS e SMITH, 1999).

As informações de custo e de cronograma relativas a um determinado produto são seus próprios custos parciais e totais, os cronogramas de projeto e de produção e dados de consumo de material, de equipamentos e de mão de obra. Elas devem ser disponibilizadas em formato numérico visando a possibilidade de cálculo de um parâmetro para a estimativa desejada. Outras informações, denominadas técnicas, que descrevem o mesmo produto e se relacionam com os custos também são necessárias. Elas correspondem às características físicas, de desempenho e de engenharia do produto. Tais itens são levantados a partir de desenhos de engenharia, especificações, certificados e experiências diretas e não necessariamente devem apresentar-se em formato numérico (OTERO, 2000).

Assim, na montagem dos bancos de dados devem ser identificados produtos similares àquele que se quer realizar o estudo de estimativa. Para isso, recolhe-se informações referentes aos custos relacionados à produção deles e a suas principais características, especialmente aquelas que mantêm larga influência sobre tais custos.

Ao se fazer uso de parâmetros disponíveis em uma estimativa, sem o conhecimento da semelhança ou diferença em relação ao produto sob estudo pode-se produzir resultados distorcidos por não se basear corretamente na realidade do projeto em questão. Isto requer que o estimador tenha bastante experiência ou se cerque da experiência de outros avaliadores e de publicações fidedignas de uso corrente (DEPARTMENT OF DEFENSE, 1999).

A montagem de um banco de dados histórico, relacionado aos produtos da própria empresa, é um importante fator para a confiabilidade de futuras estimativas de custo. É ideal que o custo apontado para seus produtos dentro de uma estimativa, seja baseado na experiência prática que a empresa tem e não em dados genéricos apresentados mediante levantamentos com menores exigências de precisão, podendo assim manter-se um maior controle sobre as variáveis que influenciam o custo (LOSSO, 1995).

Fontes externas de dados podem ser usadas contendo informações de outras empresas, dados diversos consolidados em publicações técnicas, além de arquivos públicos. Tais fontes, apesar de serem freqüentemente utilizadas, apresentam algumas desvantagens, como a falta de conhecimento do usuário a respeito dos procedimentos utilizados para levantamento de dados, a questão do tratamento estatístico dos dados originais, as diferenças entre os processos tecnológicos utilizados, além da inabilidade de se prever mudanças em futuros índices (DEPARTMENT OF DEFENSE, 1999).

2.7 Normalização de Dados

Os dados após coletados devem passar por um processo de análise e normalização. Para normalizá-los, os dados históricos são ajustados em relação a um denominador comum, os quais através desse nivelamento tornam-se disponíveis à comparabilidade. Todo o banco de dados precisa ser ajustado de certos efeitos para tornar seus elementos homogêneos ou consistentes. Para tal, deve-se examiná-los para assegurar que estejam livres dos efeitos de:

- variação do valor monetário com o passar do tempo;
- melhoria do custo devido ao ganho de produtividade;
- variação das taxas de produção durante o período em que os dados são coletados (DEPARTMENT OF DEFENSE, 1999).

Ajustes mais comuns são relativos à inflação, que é tratada como uma elevação no nível geral de preços, sem um aumento na produção ou produtividade. Para o ajuste dos dados relacionado com a inflação, são usados índices, cujas informações podem ser internas ou externas. As informações internas são a base particular de uma empresa ou de uma organização para formação de seus próprios índices. Por

outro lado, as informações externas são as publicações periódicas de índices por órgãos oficiais ou instituições que estudam as variações da economia de seu país. Os exemplos de informações externas incluem o Índice de Preços ao Consumidor (IPC) e outras previsões de inflação de vários modelos econométricos (DEPARTMENT OF DEFENSE, 1999). A relação de equivalência dólar versus Real também representa um índice usado para nivelamento dos custos com data de referência diferente (JUNGLES, 1994).

Portanto, os dados de custo devem ser ajustados em uma data de referência fixada, usando o mesmo tipo de índice de inflação para homogeneizar estes dados anterior ou posterior à data de referência e levando-os à data de estimativa a considerar.

A melhoria do custo é também um efeito que acontece ao longo do tempo, isto é, à medida que aumenta a eficiência na produção de unidades sucessivas, ocorre a melhoria do custo. Esse efeito poder ser representado pela teoria da “Curva do Aprendizado”. A teoria é adaptada para responder pela melhoria de custo em relação ao ganho de eficiência de uma empresa, e é definida pela equação seguinte:

$$Y=AX^b$$

Onde:

Y = horas/unidade (ou valor monetário/unidade);

A = horas da primeira unidade (ou valor monetário/unidade);

X = número da unidade para a qual as horas (ou valor monetário) são desejadas;

b = inclinação da curva relacionada com o aprendizado (uma constante refletindo a taxa de horas ou valor numérico que decrescem conforme a quantidade de produção dobra).

Em modelos paramétricos, a curva do aprendizado é freqüentemente usada para analisar o custo direto de unidades fabricadas sucessivamente (DEPARTMENT OF DEFENSE, 1999).

A teoria da curva de melhoria do custo teve muitas inovações desde sua origem. Uma das inovações mais comum é a introdução de uma variável na equação para incluir o efeito de variação da taxa de produção. A taxa de produção é definida como o número de itens produzidos em relação a um determinado período de tempo. A

equação adiante modifica a fórmula geral de melhoria de custo, incluindo a taxa de produção (Q'):

$$Y = AX^b \cdot Q'$$

Onde:

Y , A , X e b já são definidos anteriormente;

Q = taxa de produção (quantidade produzida durante o período de produção);

r = inclinação da curva relacionada à taxa de produção.

A razão de incluir na equação a influência de Q' é ajustar o valor monetário da primeira unidade (A) para várias taxas de produção ao longo de toda a atividade de produção. A equação também produz uma inclinação influenciada pela taxa relacionada ao aprendizado. A equação afetada pela taxa deve ser monitorada para problemas de colinearidade. Sempre que possível, os efeitos da taxa devem ser derivados de padrões de comportamento de programas de dados históricos, observados de acordo com a mudança das taxas de produção, enquanto mantêm o coeficiente de inclinação de aprendizado constante. O efeito da taxa pode variar consideravelmente dependendo do que for exigido para efetuar a mudança (DEPARTMENT OF DEFENSE, 1999).

Outros ajustes mais complexos também são usados na análise paramétrica. Dentre eles, estão os relacionados com anomalias (eventos incomuns), cujos efeitos nos custos não devem ser esperados nos projetos novos.

2.8 Análise de Direcionadores de Custo

Numa relação paramétrica, o custo estimado de um item ou atividade é função de uma ou mais variáveis relevantes independentes. Estas variáveis são os direcionadores de custo, as quais podem ser elementos de custo, tais como custos ou horas. Elas ainda podem consistir de elementos referentes às características ou às propriedades, geralmente técnicas, de um item final a estimar. Como se depreende, uma relação paramétrica pode também relacionar custo com custo (DEPARTMENT OF DEFENSE, 1999).

No desenvolvimento de relações paramétricas de custo, deve-se dar ênfase aos direcionadores de custo. Estes direcionadores são aquelas características de um

produto ou item que tem maior efeito sobre o seu custo. Um direcionador de custo pode ser uma propriedade física, química, visual, funcional ou qualquer outra identificável de um produto ou item (LONG, 2000).

2.9 Processo de Desenvolvimento de Relações Paramétricas

O processo de desenvolvimento de relações paramétricas inicia-se pela identificação de uma oportunidade para usá-la na determinação do custo de um item ou produto, visando melhorar a sua estimativa. Como resultado desse passo obtêm-se uma descrição da oportunidade específica, das necessidades de dados, das ferramentas de análise, dos critérios de aceitação de relação paramétrica e de um processo planejado para mantê-la atual. Para tal, devem ser investigadas, tipicamente, várias relações de estimativa. Entretanto, a busca de oportunidades de relações paramétricas, de vez em quando são particularmente ineficientes, principalmente quando elas precedem a coleta de dados. Então, deve-se pesquisar os bancos de dados para que eles possam apoiar as oportunidades simultaneamente com a coleta.

A qualidade de uma relação paramétrica depende da validade do banco de dados onde ela é desenvolvida, e subseqüentemente o modo como é usada em estimativas futuras. A definição da adequabilidade de uma relação paramétrica e sua aplicabilidade, no sistema que está sendo estimado, requer uma análise completa desse sistema e o conhecimento do banco de dados (DEPARTMENT OF DEFENSE, 1999). Todavia, é possível fazer algumas observações sobre o desenvolvimento de relações paramétricas de custo. Elas são equações analíticas que relacionam várias categorias de custo com seus direcionadores ou suas variáveis explicativas. Essas podem ter numerosas formas e variam de simples analogias a funções matemáticas formais obtidas da análise estatística de dados empíricos. As relações paramétricas de custo podem ser lineares ou não-lineares com uma ou diversas variáveis independentes (direcionadores de custo). Conforme elas passam de uma relação linear para não-linear, aumentam de complexidade. Conforme passam de um único direcionador de custo para mais de um, a complexidade também aumenta. A situação mais complexa é uma relação não-linear com mais de um direcionador de custo (LONG, 2000).

Independentemente do grau de complexidade, desenvolver uma relação paramétrica de custo requer um esforço combinado para coletar e refinar os dados que constituem sua base empírica. Coletar um banco de dados confiável é especialmente importante e, freqüentemente, é a atividade mais demorada. A razão número um para se produzir relações paramétricas não válidas consiste na falta de um banco de dados apropriado. Ao desenvolvê-las, deve-se primeiro formular hipóteses e então testar as relações de estimativas lógicas. Por exemplo, faz sentido supor que os custos aumentarão à medida que as exigências de propulsão do motor da aeronave aumentam? Se isso faz sentido, será necessário refinar aquela hipótese para determinar se a relação é linear ou não. Depois de desenvolvida uma relação hipotética, é necessário organizar o banco de dados para testar as relações propostas. Às vezes, ao se coletar um banco de dados, descobre-se que os dados brutos estão, pelo menos, parcialmente no formato errado para propósitos analíticos, ou mostram irregularidades e inconsistências. Ajustes nos os dados brutos, entretanto, quase sempre precisam ser feitos para assegurar um banco de dados razoavelmente consistente e comparável. É importante destacar que nenhum grau de sofisticação, no uso de estatísticas matemáticas avançadas, pode compensar um banco de dados seriamente deficiente (DEPARTMENT OF DEFENSE, 1999).

Sabendo-se que o problema de dados é fundamental, um tempo considerável é comumente dedicado à coleta de dados, ao ajuste desses dados para assegurar consistência e comparabilidade, e a adequação do armazenamento da informação para que possa ser rapidamente recuperada quando necessária.

Um esforço maior é dedicado tipicamente a coleta de um banco de dados de qualidade que qualquer outro passo no processo. Dada a informação apropriada, portanto, a tarefa analítica de produzir equações paramétricas de custo é, com freqüência relativamente fácil.

O assunto relativo a um bom banco de dados é complementado pela necessidade de primeiro fazer hipóteses e depois testar a formula matemática da relação paramétrica de custo.

Discute-se o que é prioritário. Há os que defendem que é a hipótese e depois a pesquisa dos dados para construir um bom banco de dados. Outros acreditam que

primeiro deve-se investigar os dados e sua disponibilidade, e, na seqüência a caracterização do modelo deve ocorrer. Independentemente da posição tomada, pelo menos em boa parte, elas acontecem simultaneamente, seja para verificar quais as variáveis disponíveis para desenvolver o modelo e selecionar os dados potencialmente úteis, seja para descartar aqueles considerados estranhos ou inúteis dentro do banco de dados. Portanto, deve-se definir a relação de estimativa, estruturar o modelo de previsão e testar a relação paramétrica proposta. O trabalho pode tomar várias formas dependendo das necessidades de previsão. Para isso envolve identificar as potenciais variáveis direcionadoras de custo, através da análise detalhada dos elementos de custo e técnica e de identificação das ligações entre eles (DEPARTMENT OF DEFENSE, 1999).

2.10 Modelação Matemática da Relação Paramétrica de Custo.

Uma vez que o banco de dados esteja desenvolvido e se tenha uma hipótese definida, a relação paramétrica de custo está pronta para ser modelada matematicamente.

A relação de estimativa pode variar em complexidade, a partir de algo bastante simples; como uma expressão numérica de valor, até algo mais complexo, como uma expressão matemática multi-variável.

Em ordem crescente de complexidade, são citados: fator , razão, métodos gráficos, análises de regressão linear simples, de regressão linear múltipla e de regressão não linear.

2.10.1 Fator e Razão

Como ocorre com outras aproximações paramétricas, o uso de um fator ou uma razão permite ao estimador elaborar uma grande parte de uma estimativa com dados históricos limitados. Assim, fator é uma expressão aritmética simples onde o custo é diretamente proporcional a uma única variável independente, e freqüentemente usado em áreas como experimentação, dados censitários, engenharia de sistemas , e administração de programa quando a falta de definição ou de tempo não permite a estimativa detalhada.

Por outro, razão é usada de uma maneira semelhante e inclui relações de custo recorrente com custo não recorrente e custo de produção da primeira unidade com custo do protótipo também da primeira unidade. Diante disso, fatores e razões podem ser inestimáveis ferramentas precisas de estimativa (LONG, 2000).

2.10.2 Método Gráfico

Para aplicar o método gráfico, os dados devem ser primeiramente lançados em um diagrama apropriado. O passo seguinte é a construção da curva que deve ser contínua e regular. Ela deve passar entre os pontos plotados, deixando-os, aproximadamente, em igual número de ambos os lados da linha e não apenas ligar os pontos. O objetivo é o melhor ajuste da curva para os pontos plotados, isto é, cada ponto é igualmente importante e a linha deve ser ajustada considerando cada ponto e a forma de dispersão de todos eles.

Embora essa técnica seja considerada ultrapassada, antes de desenvolver uma regra de previsão ou uma equação matemática, é aconselhável, em todos os casos, plotar os dados e observar a dispersão dos pontos e a possibilidade de qualquer ponto requerer investigação adicional (DEPARTMENT OF DEFENSE, 1999).

2.10.3 Análise de Regressão

A análise de regressão é usada para prever uma variável a partir de um conjunto de variáveis ou de apenas uma, ou seja, é a associação estatística de uma ou mais variáveis independentes com uma variável dependente.

A regressão é uma ferramenta poderosa, pois permite conhecer eventos passados, presentes ou futuros através de informações sobre eventos passados ou até presentes. Ela é empregada com frequência porque se ganha tempo e/ou dinheiro na coleta de informações para fazer previsões ao invés de coletar informações sobre o próprio evento, principalmente, quando ele ocorre no futuro (STOCKBURGER, 1996).

A análise de regressão é, matematicamente, o estudo e a prática do ajustamento de funções analíticas, tais como, $y = f(x)$ no caso unidimensional e $z = f(x,y)$ no caso bidimensional para dados que podem ser o resultado de medidas de séries temporais de uma quantidade física e real.

Basicamente, há dois tipos de análise de regressão:

- o linear, quando $f(x) = p_1f_1(x) + p_2f_2(x) + p_3f_3(x) + \dots$ para um conjunto de parâmetros $\{p(i)\}$ e um conjunto de funções $\{f(i)\}$ de x ;
- o mais geral, não linear, quando $f(x) = f(p_1, p_2, p_3, \dots, x)$, isto é, $f()$ é uma função de x e do conjunto de parâmetros de ajuste $\{p(i)\}$ (PRESS ET ALL, 1992).

No caso da análise de regressão linear, destacam-se duas formas:

- a regressão linear propriamente dita, que é a mais comum na estatística, onde uma simples variável (x) independente é usada para estimar a variável (y) dependente e a relação delas assume a equação de uma linha reta;
- a regressão múltipla que ajusta dados para um modelo que define (y) como uma função de duas ou mais variáveis (x) independentes.

2.10.3.1 Regressão Linear

A regressão linear é definida como a busca da melhor relação entre duas variáveis através do estabelecimento de uma equação matemática linear (reta).

Caso se queira prever os valores de uma variável que esteja relacionada com outra, cujos valores sejam previamente conhecidos, deve ser usada uma equação de regressão. Por exemplo, para obter a resistência de um metal, com base na sua dureza, se porventura apresentarem uma relação entre si, basta conhecer a dureza (STEVENSON, 1986). Assim, não haveria necessidade de se destruir o metal para saber sua resistência.

A regressão linear significa que a relação funcional das variáveis pode ser descrita graficamente (sobre um sistema de coordenadas X-Y) por uma linha reta e matematicamente através da equação:

$$Y = A + BX, \text{ onde :}$$

Y representa o valor estimado da variável dependente,

X é a variável independente,

B é a declividade da reta ou também coeficiente angular,

A é o ponto ou cota onde a reta intercepta o eixo vertical (Y).

A equação de regressão de duas variáveis consiste em duas partes, a parte funcional e a parte aleatória. Assim, uma população pode ser representada pela equação $Y = A + BX + E$. Onde, a porção " $A + BX$ " é a parte funcional (uma linha reta), e E (o termo do desvio) é a parte aleatória. A parte do resíduo (E) está sempre presente por causa dos erros de determinação, medida e observação. Esses tipos de erros existem com frequência devido às limitações humanas e às limitações associadas aos eventos do mundo real (DEPARTMENT OF DEFENSE, 1999).

Visto que é praticamente impossível obter os dados de uma população inteira, normalmente trabalha-se com uma amostra dessa população. Assim, trabalha-se com uma amostra para adaptar a equação de regressão para a forma: $Y = a + bX + e$. Portanto, novamente o termo " $a + bX$ " representa a parte da função da equação e " e " a parte aleatória. A estimativa dos parâmetros verdadeiros " A " e " B " da população é representada na equação da amostra por " a " e " b " respectivamente. Dessa forma, " a " e " b " são estatísticos. Isto é, eles são estimativas dos parâmetros da população. Como parâmetros estatísticos, eles estão sujeitos aos erros de amostragem. Conseqüentemente, um bom planejamento de amostras aleatórias é importante (Department of Defense, 1999).

A aproximação matemática denominada método dos mínimos quadrados é a raiz da análise de regressão, que pode ser definida como a natureza matemática da associação entre duas variáveis. Essa técnica, usada para melhorar a habilidade de prever o valor mais próximo do real da variável dependente, especifica uma linha que melhor ajuste o conjunto de dados. O método faz isso minimizando a soma dos quadrados dos desvios que resultam da diferença entre os valores observados de Y e os valores calculados de Y . O valor observado representa o valor realmente existente na base de dados e o valor calculado, identificado como Y_c , é aquele valor estimado pela equação tomando o mesmo valor de X . A diferença entre os valores observado e calculado é representada pelo termo " e " (desvio) da equação linear (DEPARTMENT OF DEFENSE, 1999).

A técnica dos mínimos quadrados analisa cada par (X, Y) na base de dados, refinando os parâmetros para os termos da declividade e do intercepto vertical até

encontrar uma equação para a linha que minimiza a soma dos quadrados dos desvios.

Para calcular a linha de melhor ajuste dos mínimos quadrados para uma base de dados de n números de observações, é necessário encontrar a função $(a + bX)$ que minimize:

$$\min \sum_{i=1}^n e_i^2 = (Y_1 - Y_{C1})^2 + (Y_2 - Y_{C2})^2 + (Y_3 - Y_{C3})^2 + \dots + (Y_n - Y_{Cn})^2$$

Os valores de a e b para a reta $Y_c = a + bX$ que minimiza a soma dos quadrados dos desvios são obtidos da solução de um sistema de equações normais (STEVENSON, 1981):

$$(1) SY = na + b(SX)$$

$$(2) SXY = a(SX) + b(SX^2)$$

Referências contidas em muitos livros de estatística mostram que essas duas equações satisfazem as exigências da regressão comum dos mínimos quadrados.

As propriedades da técnica são:

- ela considera todos os pontos de dados;
- a soma dos quadrados dos desvios entre a linha e os pares de dados observados é o valor mínimo possível.

Similaridades entre estas duas propriedades e a média aritmética são observadas. A média aritmética é a soma dos valores da variável independente divididos pelo número de observações e a soma dos valores da variável dependente dividida pelo número de observações. Isso resulta que o ponto (\bar{Y}, \bar{X}) relativo às médias cai sobre a reta dos mínimos quadrados.

Das equações normais são derivadas duas equações para calcular a e b diretamente :

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$a = \bar{Y} - b\bar{X}.$$

(Lembrando que uma vez conhecido b , é possível resolver a equação da reta $Y = a + b X$ para a , porque é sabido que o ponto (\bar{Y}, \bar{X}) deve cair sobre a reta.)(DEPARTMENT OF DEFENSE, 1999).

Então, obtidos os valores de a e b , a equação da reta que melhor ajusta uma base de dados está definida.

Note-se que a regressão linear não testa se seus dados são lineares (exceto via “runs test”). Ela assume que seus dados são lineares e busca os coeficientes angular b e o linear a que produzam uma linha reta que melhor ajuste seus dados

Logo, segundo (MOTULSKY, 1999), quando se trabalha com a análise de regressão linear tem-se a considerar um certo número de limitações, erros e cuidados, que a seguir serão analisados.

- Em muitas pesquisas a relação entre X e Y é uma linha curva e, portanto, produz a regressão linear inadequada. Ou os dados podem ser transformados ou então usa-se o ajuste da regressão não linear.
- A análise de regressão linear assume que a dispersão dos pares de dados ao redor da linha é Gaussiana.
- A variância de cada conjunto da variável independente é constante, portanto, a hipótese é violada se os pontos com valores altos ou baixos de X tenderem a ficar mais distante da linha de melhor ajuste.
- Se um ponto está acima ou abaixo da linha é um mero acaso, e não influencia se outro ponto está em quaisquer dessas posições. Quando a hipótese é violada, os dados são autocorrelacionados. Essa hipótese requer que o termo do desvio seja uma variável verdadeiramente aleatória.
- O valor médio de cada distribuição cai sobre a linha de regressão.
- O modelo de regressão especifica que a variável independente deve ser um número fixo e não uma variável aleatória.
- Uma verificação lógica deve ser feita, determinando uma hipótese apropriada e analisando a base de dados de forma que uma avaliação possa ser relacionada a causa e efeito. Por exemplo, assumir um alto grau de correlação entre o número de telefones públicos e a venda de bebidas numa cidade.

Claramente conclusivo que não há causa e efeito envolvido no problema. Uma variável mais lógica, como a população da cidade, torna-se uma variável independente de maior causa que direciona tanto o número de telefones quanto à venda de bebidas.

- Se a população base não é mais relevante devido às mudanças na tecnologia, por exemplo, então a equação de regressão linear pode não ser a melhor ferramenta de previsão para se usar. Ao se usar um modelo de regressão deve-se ter certeza que os dados não se tornaram obsoletos para previsão no momento atual ou no futuro.

2.10.3.2 Regressão Múltipla

Basicamente a regressão múltipla é a associação de três ou mais variáveis, ou seja, nada mais que uma extensão da regressão linear simples. Nesse caso, há apenas uma variável dependente, e tem-se duas ou mais variáveis independentes (ditas também explanatórias).

A análise de regressão múltipla assume que as variações em y podem ser melhor explicadas associando-as a mais de uma variável independente.

A finalidade das variáveis independentes adicionais é melhorar a capacidade de predição se comparada com a regressão linear simples (STEVENSON, 1981).

A equação de regressão múltipla pode exprimir y em função de variáveis independentes x_1, x_2, \dots e ser denominada equação de regressão de y para x_1, x_2, \dots

Por exemplo, no caso de três variáveis independentes, a equação tem a seguinte forma:

$$Y_c = a + b_1x_1 + b_2x_2 + b_3x_3,$$

em que a, b_1, b_2, b_3 são constantes, sendo:

y_c = o valor estimado da variável dependente;

a = o intercepto vertical, quando $x_1, x_2, x_3 = 0$;

x_1 = a primeira variável explanatória;

b_1 = o coeficiente angular da linha relativo às variações em x_1 ;

x_2 = a segunda variável explanatória;

b_2 = o coeficiente angular da linha relativo às variações em x_2 ;

x_3 = a terceira variável explanatória;

b_3 = o coeficiente angular da linha relativo às variações em x_3 (DEPARTMENT OF DEFENSE,1999) (STEVENSON, 1981)

Encontrar as combinações das variáveis explanatórias não é tarefa fácil. Em alguns casos, os resultados de regressão múltipla podem parecer paradoxais. O modelo ajusta bem os dados, embora nenhuma das variáveis independentes tenham estatisticamente um impacto significativo na predição de y . Quando duas variáveis independentes apresentam alta correlação, ambas carregam essencialmente a mesma informação. Nesse caso, nenhuma pode contribuir significativamente ao modelo depois que a outra é incluída, entretanto, juntas elas contribuem muito. Removendo ambas as variáveis do modelo o ajuste seria muito pior. Assim, o modelo global ajusta bem os dados, mas nenhuma variável x faz uma contribuição significativa quando é adicionada ao seu modelo final. Quando isso acontece, as variáveis independentes são colineares e os resultados mostram multicolinearidade (DEPARTMENT OF DEFENSE,1999).

Se o objetivo é simplesmente predizer y de um conjunto de variáveis x , então a multicolinearidade não causa problema. As predições ainda serão precisas.

2.11 Teste de Significância da Relação Paramétrica de Custo

Depois de discutir a técnica de regressão , volta-se para avaliar a qualidade da estimativa paramétrica de custos. Assim deve-se responder as perguntas, conforme DEPARTMENT OF DEFENSE (1999):

- é significativamente boa uma equação de estimativa paramétrica de custo e está significativamente boa para estimar o custo de itens específicos ou serviços?
- qual é o intervalo de confiança da estimativa (i.e., como provavelmente é o custo estimado para se cair dentro de uma faixa especifica de resultados de custo)?

Muitos analistas contam com duas estatísticas primárias para fazer essa determinação: o coeficiente de correlação (R) e o coeficiente de determinação (R^2). Ambas as medidas indicam simplesmente o grau de relacionamento entre as variáveis.

Nenhuma medida indica causa e efeito. Causa e efeito requerem uma verificação de lógica e dependem da perspicácia do analista.

Há várias outras estatísticas para avaliar e para ampliar o conhecimento e a confiança na equação de regressão assim como se assegurar de sua capacidade de previsão.

Para avaliar a qualidade de uma relação paramétrica de custo, há necessidade de uma relação de itens para examinar. Há elementos que focam na validação estatística da relação paramétrica ou modelo, tais como: a qualidade dos dados, a relação lógica entre eles, os testes estatísticos t e F , o desvio padrão, o coeficiente de variância e o coeficiente de determinação (R^2). Enquanto outros, como o número de observações, grau de liberdade, “outliers” e amplitude dos dados, enfocam no uso da relação ou modelo para predizer estimativas futuras.

É justificada a precaução quando se executa a análise estatística de uma relação paramétrica de custo. Não há nenhuma estatística que desqualifique uma relação paramétrica de custo ou um modelo nem há qualquer estatística que “valide-os”.

A matemática que modela o esforço de estimativa deve ser examinada à luz de uma perspectiva completa e começar com os dados e a lógica da relação. Assim, é importante uma completa explanação narrativa da qualidade do banco de dados e da lógica do modelo proposto. Só depois de assegurar que os dados e a lógica da relação são sólidos é que se deve começar a avaliar a qualidade estatística do modelo.

O exame estatístico começa tipicamente com uma avaliação das variáveis individuais no modelo. O método mais comum para avaliar a significância da variável na relação é o teste t para cada variável explanatória. O próximo passo a avaliar é a significância da equação inteira. O teste F é a estatística mais comum para avaliar esta qualidade. Ao assumir que as variáveis e a equação inteira têm significância, o passo seguinte é julgar o tamanho e a proporção do desvio estimado da equação. O desvio padrão da estimativa (DPE ou DP) e o coeficiente de variância (CV) fornecem essa compreensão.

Finalmente, a análise estatística típica conclui com o exame do valor do coeficiente de determinação (R^2) ou R^2 ajustado ao comparar os modelos com um número diferente de variáveis independentes para cada modelo. O coeficiente de

determinação mede a porcentagem de variação da variável dependente explanada pelas variáveis independentes.

Os elementos usados na predição de estimativas futuras enfocam na geografia dos dados, a partir da qual a relação paramétrica de custo ou o modelo é construído. O ideal preferível é um modelo estatístico forte com um grande número de observações e com um menor número de variáveis para formular a equação. Além disso, deve haver um pequeno número de pontos de dados reais que o modelo prediz pobremente. Conclui-se assim que a parte crítica de qualquer avaliação é identificar a amplitude dos valores independentes nos quais o modelo é construído. Teoricamente, o modelo só é válido dentro dessa amplitude relevante dos dados de valores independentes. Na prática, enquanto a relação matemática assumida permaneça válida, o uso do modelo é admissível fora dessa faixa. É provável que isso seja só um pequeno limite além dos valores atuais dos dados. A faixa de validade estabelece sobre que amplitude a relação paramétrica fornecerá predições razoáveis.

Não há padrão definido relacionado ao critério de aceitação para as várias estatísticas. Não há nenhum umbral absoluto. A análise estatística; o processo de normalização, a coleta de dados do analista e a lógica associada formam toda a base para aceitar a relação paramétrica como a base para se estimar.

A seguir é apresentada a interpretação não estatística dos diversos indicadores estatísticos anteriormente referenciados.

- teste F: testa se a equação inteira, como um todo, é válida;
- teste t: testa se as variáveis X individuais são válidas;
- desvio padrão (DP): média que calcula o desvio quando usar a equação como a regra de estimativa;
- coeficiente de Variância (CV): desvio padrão dividido pela média dos dados de Y, medida relativa do desvio estimado;
- coeficiente de determinação (R^2): porcentagem da variação nos dados de Y explanados pelos dados de X;
- $-R^2$ ajustado: coeficiente de determinação ajustado para o número de variáveis X usadas para explicar a variação nos dados de Y;

- graus de liberdade (GL.): número de observações (N) menos o número de parâmetros estimados (variáveis de $X + 1$ para o termo constante “a”). Conceito de parcimônia aplicado com isso num modelo preferido é aquele com alta significância estatística que usa o menor número de variáveis;
- “outliers”: observações de Y que o modelo prediz com pobreza. Isso não é sempre uma razão válida para descartar os dados;
- valor P: nível de probabilidade no qual o teste estatístico falharia e sugeriria que a relação não seria válida. Valores P menores que 0,10 são geralmente preferidos (i.e., somente uma chance de 10% ou menos de que o modelo não é bom) (DEPARTMENT OF DEFENSE, 1999).

2.12 Validação do Modelo Paramétrico.

Para determinar se uma relação paramétrica é boa, para predição dos custos, é necessário avaliar sua precisão. O processo ou ação, utilizada para demonstrar essa precisão, é definida como validação. A validação também inclui uma demonstração que as relações de dados são lógicas e fortes, os dados usados são confiáveis e que as relações paramétricas predizem com precisão, baseadas na história passada, a experiência atual ou futura.

A melhor técnica de validação é a usa dados independentes do conjunto de dados históricos que geraram o modelo paramétrico. Por exemplo, se uma companhia tem dados históricos suficientes, podem ser separados pares de dados independentes do processo de construção do modelo e usados como elementos de teste para avaliar a precisão. Infelizmente, devido às restrições de dados isso nem sempre é possível nem é a melhor prática. Frequentemente, os conjuntos de dados disponíveis são extremamente pequenos. Ao separar alguns pares de dados do desenvolvimento do modelo para validação pode-se omitir informação que poderia ser usada para desenvolver os parâmetros de um modelo mais preciso. Esse aspecto precisa ser considerado durante o desenvolvimento do modelo (DEPARTMENT OF DEFENSE, 1999).

Quando os dados históricos são suficientes, mas não estão disponíveis para propósitos de prova independentes, podem ser feitas avaliações de precisão usando outras técnicas. Assim, uma comparação pode ser feita entre a estimativa

desenvolvida através do modelo construído, e uma preparada com outras técnicas de estimativa.

É necessário usar de bom senso quando da determinação do nível de precisão, pois não há uma definição do nível de precisão do modelo da relação paramétrica. Em geral, as relações paramétricas devem ter tanta acuracidade quanto à técnica de estimativa geradora da fonte de dados (DEPARTMENT OF DEFENSE, 1999).

2.13 .Modelos Paramétricos de Custos Aplicados na Construção e no Saneamento.

PANZETER (1993) investiga em sua pesquisa projetos de pontes e analisa uma metodologia alternativa de processos que otimiza estimativas de custo e uma melhor estratégia de estimativa de tempo de duração da construção.

São analisados três tipos de pontes de concreto de rodovia utilizando um modelo dividido em cinco grupos ou categorias de serviços. O custo total do projeto é a soma dos custos de todas as categorias. Os grupos são denominados em: Infra-estrutura, Superestrutura, Encontros, Tráfego e Outros. Cada categoria tem seu custo modelado pela aplicação de técnicas estatísticas de regressão múltipla.

Dois tipos de modelos são considerados na análise estatística das relações paramétricas de custo: o “geométrico”, que usa como direcionadores as características físicas (ex.: comprimento da ponte) e “materiais”, cujos direcionadores são as quantidades dos principais materiais (ex.: volume de concreto).

O modelo de melhor ajuste, isto é, o mais apropriado de acordo com a estatística é o considerado. No caso da categoria Infra-estrutura, o modelo “materiais” é o mais adequado, enquanto que o “geométrico” é o mais apropriado nas categorias Superestrutura e Outros. A combinação dos dois modelos utilizados gera um terceiro, que é o escolhido para explicar os custos da categoria Encontros. O custo da categoria Tráfego não encontra relação com as dimensões e nem com as quantidades de materiais, por isso é considerado um valor constante.

Observa-se, no modelo da categoria Superestrutura que os tipos de ponte considerados (Laje, Viga I e Viga Caixão) não representam variáveis significantes.

O mesmo acontece com a variável tipo de estrada. Na categoria Outros são descartados seis pares de dados considerados muito afastados dos demais pontos (outliers).

Os modelos são testados para nove projetos e seus resultados comparados com os custos de concorrência pública e com os custos reais. A comparação mostra que os modelos estão a apenas 7% do custo de licitação e 6% do custo real. Isso demonstra que a precisão dos modelos é altamente significativa.

OTERO (2000), afirma que ao estudar a aplicação de relações paramétricas para estimativa de custo de construção de edifícios, usa-se como direcionadores de custo as características gerais das obras, disponíveis nas etapas iniciais do processo construtivo. A área total de construção, a área do pavimento tipo, a área do pavimento de subsolo, o número total de banheiros, o número de elevadores e o tempo de construção são as características utilizadas.

O autor citado ainda estabelece relações entre variáveis ligadas a materiais, a mão-de-obra e a serviços e aos direcionadores que lhe parecem de maior afinidade. Tais relações são, então, selecionadas e avaliadas a partir de razões e de regressões lineares que envolvem esses elementos citados apoiando-se especialmente nos indicadores relativos a variabilidade.

Os modelos paramétricos produzidos fazem uso daquelas regressões para explicar os custos a partir de seus direcionadores a cada um dos grupos de serviços. Para isso, escolhem-se aquelas equações que determinam uma maior confiabilidade na predição dos custos. Construídos esses modelos, é feita a comparação entre a precisão que eles alcançam na estimativa dos custos e aquele valor estabelecido pela simples razão entre o custo global e a área total do edifício.

A área do pavimento tipo é a variável de maior correlação com a área total do edifício, tendo a regressão linear alcançado um coeficiente de determinação ajustado de 97,40%.

Com resultados pouco inferiores em termos de associação, a área do subsolo também apresenta um alto nível de correlação com a área total do edifício.

Enquanto a variável tempo de construção não apresenta boa correlação com a área total, o número de elevadores e o número de banheiros são bem correlacionados com essa área.

U.S. ENVIRONMENTAL PROTECTION AGENCY (1981) faz um estudo de custos de construção de 777 sistemas municipais de coleta de esgoto. Para realizá-lo, o sistema de coleta é desagregado em componentes, como redes coletoras e interceptores, estações elevatórias e emissários.

Os dados de custo são obtidos de licitações das obras dos sistemas de esgoto. Desses documentos, os custos de construção são separados para cada componente do sistema.

Sua base de dados inclui tipo, tamanho e profundidade de tubulações de redes de esgoto e emissários, assim como o número e capacidade de estações elevatórias e todos os custos de construção.

Os dados de custos são identificados conforme as construções ocorram na área rural ou em cidades e também pela média nacional dos E.U.A.

O componente rede coletora tem seus custos totais de construção subdivididos em custos parciais como:

- custos da tubulação na vala, explicados pela variável comprimento do tubo, são definidos pelo custo dos materiais e da mão-de-obra para escavação e assentamento e mais o reaterro da tubulação, por tipo de tubo, profundidade de assentamento e pelo local da obra (cidade ou área rural);
- custos não relacionados diretamente com a tubulação, uma porcentagem do custo anterior, são obtidos com base nas seguintes categorias: elementos acessórios (válvulas, comportas), travessias, escavação em rocha, substituição de pavimentos, interligação de redes, berços e reaterros especiais, mobilização e miscelâneas, poços de visita e tampões.

Para o emissário, os custos totais de construção são obtidos da mesma maneira que para a rede coletora.

Por outro lado, as estações elevatórias, tem o custo de construção relacionado com a capacidade de vazão de projeto (Q_d), através de três equações paramétricas, sendo uma para cada faixa de altura manométrica total (H_m), conforme a Tabela 2.1, a seguir.

Tabela 2.1. Equações Paramétricas de custo para estação elevatória

Altura Manométrica	Equação Original Q_d (galões/min)	Equação Adaptada Q_d (l/s)
$H_m > 80$ pés ($H_m > 24,4$ m.c.a.)	$\text{Custo}_a = 1,59 \times 10^3 Q_d^{0,59}$	$\text{Custo}_a = 5,10 \times 10^3 Q_d^{0,59}$
$40 < H_m < 80$ pés ($12,2 < H_m < 24,4$ m.c.a.)	$\text{Custo}_b = 1,13 \times 10^3 Q_d^{0,70}$	$\text{Custo}_b = 7,82 \times 10^3 Q_d^{0,70}$
$H_m < 40$ pés ($H_m < 12,2$ m.c.a.)	$\text{Custo}_c = 1,26 \times 10^3 Q_d^{0,75}$	$\text{Custo}_c = 10,0 \times 10^3 Q_d^{0,75}$

Fonte: Adaptado de U.S. Environmental Protection Agency (1981)

Os custos não relacionados diretamente com a construção das estações elevatórias são pequenos e não representam mais que 5% do custo de construção total delas.

Os resultados obtidos da análise da base de dados são apresentados dentro da variação normal esperada. Alguns destaques de vários itens, embora tenham dados suficientes para assegurar sua precisão, apresentam aparência anômala. Esses destaques são os seguintes:

- tubos de concreto são os mais comumente usados, seguidos de perto por manilhas cerâmicas e os tubos menores de 16 polegadas são os mais prevalentes;
- ABS é o tipo de tubo menos usado;
- são notadas pequenas diferenças de custos quando comparados os tubos assentados em profundidade menor que 8 pés com aqueles enterrados entre 8 e 15 pés;
- os tubos assentados a profundidades maiores que 15 pés de profundidade são significativamente mais caros que aqueles enterrados entre 8 e 15 pés;
- os custos de transporte ao local de construção dos tubos de 21 polegadas são mais caros quando realizados em cidades do que para os efetuados em área rural e essa tendência se reverte para os diâmetros maiores;
- o custo da mão-de-obra para rede coletora dentro de cidades é significativamente mais caro que para a área rural;

- os custos de construção de poços de visita são mais caros para as profundidades de 8 a 15 pés do que para as menores de 8 pés , isso dentro da área metropolitana e verificado na média nacional;
- poços de visita de 72 polegadas de diâmetro resultam mais caros na área rural, provavelmente porque esse diâmetro é usado apenas em situações especiais na maioria das zonas rurais;
- os custos dos tubos na vala, para emissários menores que 21 polegadas, resultam significativamente mais baratos que os equivalentes para rede coletora;
- a profundidade de assentamento de tubos de esgoto é, em média, maior em cidades que na área rural;
- foram analisados mais de 15000 itens de licitação para rede coletora, certamente a mais importante análise empírica dos custos de construção de rede disponível nos E.U.A até 1980;
- foram analisados projetos para comunidades menores que 2500 habitantes até um milhão de habitantes.

U.S. ENVIRONMENTAL PROTECTION AGENCY (1980), em seu estudo sobre Custos de Construção de Estações Municipais de Tratamento de Esgoto (1973-1978), estabelece relações paramétricas de custo de estações completas de tratamento, de processos unitários e de componentes de estações ou processos. Essas relações resultam da análise de uma base de dados de 737 projetos licitados para construção. Entre os projetos, 469 envolvem estações com tratamento secundário, 111 de tratamento secundário avançado e 157 de tratamento avançado de esgoto.

Os dados de documentos de licitações variam consideravelmente devido a localização, ao tamanho e ao tipo de projeto. Eles são analisados para todos os tipos de plantas, processos e componentes coletados. Para aqueles que a análise estatística indica curvas resultantes não confiáveis dentro de determinado nível, elas são descartadas do estudo.

Os custos de construção são analisados e desenvolvidos para três níveis de detalhe. O mais geral, denominado Custos de Primeira Ordem, são para estações completas de tratamento, onde todos os custos de construção são incluídos. O segundo nível de detalhe, Custos de Segunda Ordem, são para processos unitários

específicos. O último nível, Custos de Terceira Ordem, é definido para os custos de vários componentes requeridos, como: escavação, instalações elétricas, instrumentação, etc.

As relações de custo de construção obtidas através de regressão linear, explanadas pela vazão de projeto ou outros parâmetros, são determinadas para cada nível de dado.

Somente aquelas relações paramétricas com um Coeficiente de Correlação da amostra maior ou igual a 0,70 são consideradas. Um valor R de 0,70, para uma curva de custo indica através do Coeficiente de Determinação R^2 que uma variação de 49% no custo é explicada por variações na vazão do projeto. Todos os gráficos de curvas de Primeira e Segunda Ordem do supracitado estudo tem valor de seu R explicitado. São também apresentados os valores do teste F estatístico, pois ele testa a hipótese do Coeficiente de Correlação da amostra igual a zero contra a alternativa de que a equação como um todo define uma relação significativa entre as duas variáveis – no caso, custo versus vazão de projeto.

Esse estudo é certamente uma das mais completas análises empíricas de custos de construção desenvolvidas para estações de tratamento de esgoto. Pode ser usado para aplicações em engenharia e estimativas preliminares de custos de construção de processos unitários ou estações completas, embora não substitua processos de estimativa detalhada de projetos. Os seus resultados são médias estatísticas de um país e não necessariamente refletem as condições locais específicas de um projeto que podem mudar drasticamente os custos finais. Os custos de mão-de-obra e materiais locais são normalizados. A normalização e atualização dos custos de construção de projetos são realizados através de índices para o tempo e para o local de sua implantação.

HANSEN, GUNERMAN e CULP (1979) em um estudo de quatro volumes apresentam curvas de custo de construção, operação e manutenção para 99 processos unitários de tratamento de água. São criadas curvas para 72 processos unitários de grandes sistemas de abastecimento de água, desde 43,1 l/s até 8760 l/s e outras para pequenos sistemas de 0,11 l/s a 43,0 l/s.

Essa separação é necessária porque muitos processos aplicáveis para uma faixa de vazão não são para outra faixa, isso é comprovado, quando um processo é

aplicado em ambas as faixas que a concepção do projeto dos componentes varia significativamente. Assim a economia relativa ao tratamento de grandes vazões, freqüentemente apresenta sensíveis mudanças na inclinação das curvas de custo, principalmente entre 43,1 l/s e 215,5 l/s.

As curvas de custo de construção são desenvolvidas usando dados de custo de equipamentos fornecidos por fabricantes, dados de custo reais de construção de plantas, de partes de unidades de projetos conceituais e de dados publicados. Quando as partes de custos de unidades são usadas, para determinar custos reais de projetos conceituais, com freqüência, são utilizadas técnicas de estimativa de Padrões de Estimativa de Construção de Plantas de Processo desenvolvidos por Serviços de Engenharia Richardsons, Dados de Custo de Construção de Edificações Mean's Company e Guia Dodge para Custos de Construção de Obras Públicas. As curvas de custo, após serem produzidas, são examinadas e ratificadas por uma outra firma de consultoria de engenharia, Zurheide-Hermann Inc, ao usar técnicas de estimativa de custo similares àquelas utilizadas por empreiteiros em geral na preparação de suas propostas de licitação. Ajustes das curvas são necessários para refletir, em cada caso, as condições específicas locais, as condições geográficas ou as necessidades de energia disponíveis.

As curvas são particularmente úteis para estimar a economia relativa de alternativas de sistemas de tratamento e para uma avaliação preliminar de nível de custo a ser esperado para um projeto proposto.

Os custos de construção são desenvolvidos pela determinação e então agregação do custo de oito componentes principais: (1)Escavação e Trabalhos Locais, (2)Equipamentos, (3)Concreto, (4)Aço, (5)Mão-de-obra, (6)Tubulações e Válvulas, (7)Equipamentos Elétricos e Instrumentação e (8)Prédio ou Edificação da Unidade. Essas oito categorias são utilizadas inicialmente para facilitar a atualização precisa do custo. A divisão também é útil para ajustar os custos para o local específico, para as condições geográficas e outras condições especiais.

Os custos de construção, para cada processo unitário, são apresentados como uma função do parâmetro de projeto que melhor os explica. Por exemplo a curva do custo de construção para plantas de filtros por gravidade é traçada em função da capacidade de vazão em galões por minuto ou ainda os custos do sistema de

geração de ozônio são relacionados com a capacidade de alimentação em libras por dia. O uso de tais parâmetros de projeto permite que curvas sejam utilizadas com maior flexibilidade do que se todos os custos fossem plotados em relação a vazão.

BENJES (1976) propõe, em seu estudo, o desenvolvimento de informações de custo de construção e das atividades de operação e manutenção de processos unitários aplicáveis ao tratamento da vazão excedente de tempo seco de esgoto combinado e das unidades de estocagem.

O custo de construção é relacionado com um simples parâmetro que é associado ao tamanho da unidade. Em geral, o tamanho das unidades de tratamento corresponde à vazão de projeto.

O caminho usado, para estimar custos para unidades de tratamento, segue os passos abaixo:

- definir as unidades pela dimensão, material de construção, equipamentos, tubulações e acessórios e conceber um anteprojeto de cada unidade definindo paredes, dimensões completas e as solicitações estruturais;
- estimar quantidades dos componentes principais de custo aplicando regras simples para produzi-las, por exemplo, paredes de concreto – mínimo de 8 polegadas ou 1 polegadas por pé de altura; sapatas de concreto – dois terços do volume da parede de concreto;
- estimar custos de implantação dos principais componentes de custo: concreto, equipamento, tubulações e válvulas, escavações e edificação;
- adicionar 10 a 20% da soma dos componentes de custo para miscelâneas;
- adicionar 5 a 20% para necessidades de eletricidade e instrumentação.

Os itens assim determinados com um apropriado fator de contingência representam o custo de construção estimado do processo unitário de tratamento. Por outro lado, ao juntar os outros custos de função da planta e do processo unitário e os custos de serviços locais (tais como: interligação de tubulação, vias de acesso, passeios e ajardinamentos) é produzido o custo total de construção da estação de tratamento.

Para normalização dos dados de custo devido à inflação são utilizados índices de custo específicos, os quais são criados por U.S. Environmental Protection Agency.

São desenvolvidas diversas estimativas de custo de construção de processos unitários e para cada um deles são produzidas estimativas de acordo com a variação do parâmetro direcionador do custo. Com base na tabela de custos de construção variando com o parâmetro definidor, são desenvolvidas curvas de custos para cada processo unitário.

As estimativas de custo e respectivas curvas de custos apresentadas no citado estudo representam as condições médias e como tais servem como diretriz. As condições locais que afetam os custos são muito significantes nas considerações de estimativa de custo e isso tende a invalidar o uso destes dados em alguns projetos. afetam também os custos, as mudanças nas normas concernentes às práticas de construção e às necessidades de segurança, na tecnologia e no equipamento da planta de tratamento.

A estimativa de custo total para uma estação de tratamento é desenvolvida, então, utilizando-se a relação das funções de processo, a seleção do custo apropriado dos dados apresentados no estudo e a adição dos custos relativos aos serviços locais, dos custos relativos aos juros de capital durante a construção e dos relativos aos serviços de engenharia, administrativos e financeiros.

JUNGLES (1994) mostra em sua pesquisa, que são estabelecidas relações paramétricas para explicar os custos de construção e de operação de componentes de sistema de abastecimento de água. Os seus modelos paramétricos para componentes ou partes do sistema são todos equações de regressão não linear.

Para adutoras de água, os custos de construção são criados em dois modelos. Um para representar tubulações em materiais plásticos (PVC) e outro para tubulações em ferro fundido. Ambos são explicados por dois direcionadores: a capacidade de adução e o comprimento da adutora. Os modelos bem ajustados apresentam, respectivamente, Coeficientes de Determinação (R^2) iguais a 89% e 98%.

O custo de construção do componente estação de recalque é também bem explanado por duas variáveis independentes, a capacidade de vazão projetada e a altura manométrica. O bom ajuste da curva de regressão apresenta um R^2 de 97%.

No caso da estação de tratamento de água, que é outro componente, o seu modelo apresenta o custo associado a um único direcionador, a capacidade de vazão nominal instalada. Seu R^2 igual a 97%, demonstra o bom ajuste do modelo.

O reservatório do tipo apoiado apresenta também um modelo, cuja equação paramétrica de custo relaciona-se a apenas uma variável independente, o volume de reservação. O seu Coeficiente de Determinação é de 99%, o que significa que as variações de custo são amplamente explicadas pelas variações do volume de reservação.

O modelo de custo de construção de rede de distribuição de água é explanado por dois direcionadores. Um é o comprimento total da rede e o outro é a vazão máxima horária. É uma relação paramétrica de custo muito bem ajustada, com R^2 igual a 99%.

Todos os modelos de custo do sistema de abastecimento de água são analisados estatisticamente para um intervalo de confiança de 95%. A análise estatística é feita para os testes “t” e “F” e suas hipóteses, análise de resíduos e análise da variância. Os resíduos analisados não detectam deficiências nos modelos e os testes de hipóteses demonstram bons níveis de significância e rejeição da hipótese nula.

3 METODOLOGIA

3.1 Caracterização da Pesquisa

A pesquisa é classificada de acordo com a sua natureza, sua forma de abordagem do problema, seus objetivos e procedimentos técnicos, como uma pesquisa aplicada, quantitativa, exploratória e documental.

3.1.1 Pesquisa Aplicada

O objetivo desta pesquisa é gerar modelos de estimativa de custo para partes de sistema de esgoto, cujo universo são os projetos arquivados na Companhia Catarinense de Águas e Saneamento, assim pretende-se atender aos interesses de pesquisadores, planejadores, projetistas e organismos financiadores do saneamento do Brasil e principalmente de Santa Catarina.

3.1.2 Pesquisa Quantitativa

Os dados coletados e os modelos são quantificáveis, classificáveis e são submetidos a análises estatísticas.

Trata-se de uma pesquisa quantitativa, pois constrói hipóteses relacionadas aos modelos, envolve levantamento bibliográfico que dá mbasamento científico e analisa exemplos que estimulam a compreensão do uso de modelos de estimativa paramétrica de custo.

3.1.3 Pesquisa Documental e Experimental

Nesse aspecto, a pesquisa é elaborada a partir de dados de projetos de sistema de esgoto não tratados analiticamente, tem objetivos baseados na seleção de variáveis que expliquem os custos dos modelos e definam as formas de controle e de observação dos efeitos que essas variáveis produzam nos modelos de custo.

3.2 Fluxograma do Processo de Desenvolvimento de Modelos Paramétricos de Custo

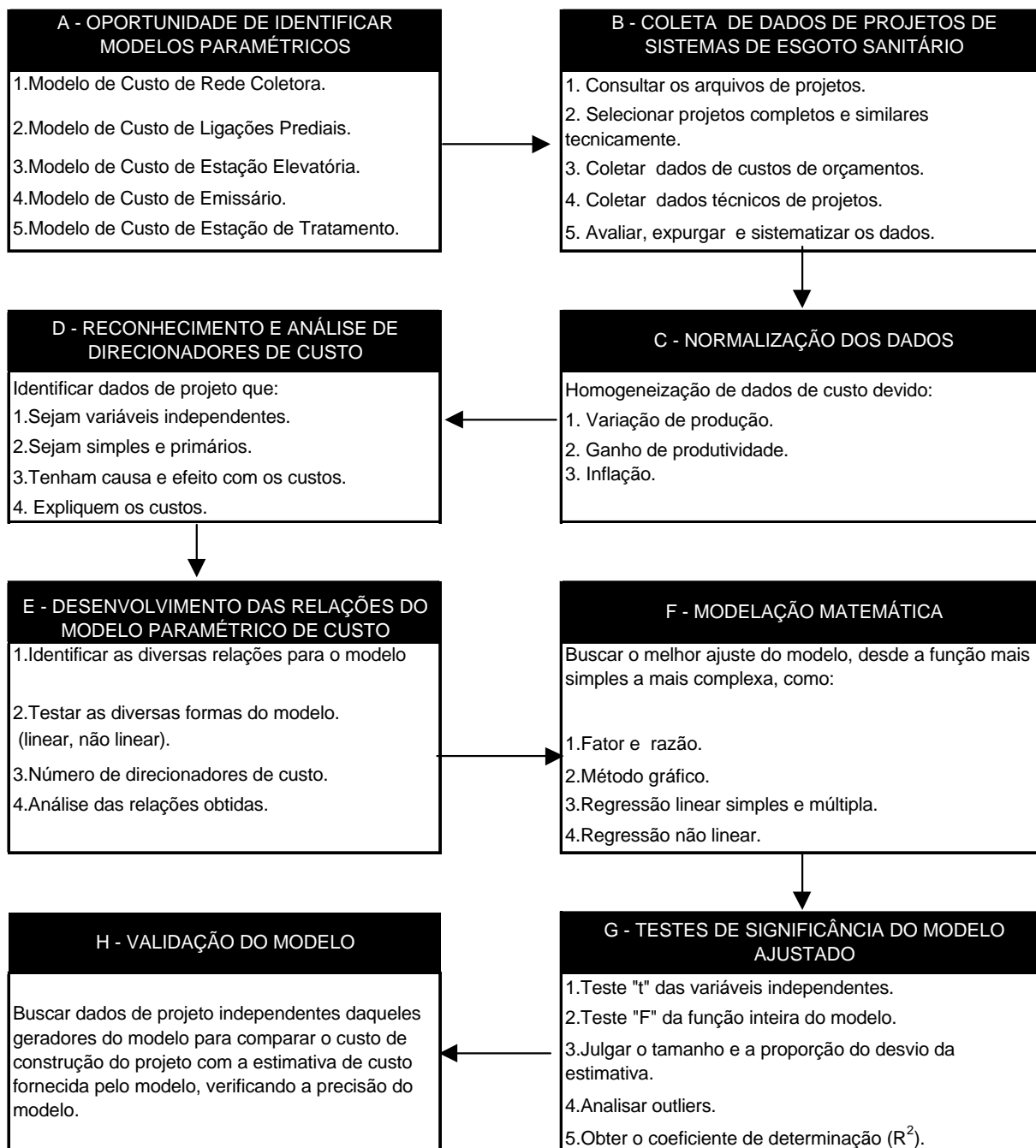


Figura 3.1. Fluxograma de desenvolvimento de modelo paramétrico de custo.

3.2.1 (A) – Oportunidade de Identificar Modelos.

O começo do processo de desenvolvimento de relações paramétricas é a identificação de uma oportunidade para usá-la na estimativa do custo de construção de projetos. Para tal, são investigadas várias relações de estimativa que possam explicar os custos de projeto de um sistema de esgoto sanitário, através do modelo de custo de:

- rede coletora;
- ligações prediais;
- estação elevatória;
- emissário;
- estação de tratamento.

3.2.2 (B) - Coleta de Dados de Projetos de Sistemas de Esgoto Sanitário.

As fontes dos dados necessários para modelação e validação das relações paramétricas de custo das partes do Sistema de Esgoto são projetos arquivados na Companhia Catarinense de Águas e Saneamento (CASAN). Os dados expressam os padrões tecnológicos adotados por essa empresa e que pouco diferem dos padrões usados em todo Brasil.

Embora essa companhia não tenha ainda muitos sistemas de esgoto em funcionamento ou projetados, cerca de 50 projetos são consultados e deles são selecionados aqueles que apresentavam-se completos e executáveis, de maiores semelhanças técnicas e de diversos tamanhos e capacidades. Desse modo, procurou-se projetos de pequenas redes coletoras com apenas o diâmetro mínimo e de outras maiores que atingissem vários diâmetros de tubulações e vários quilômetros de extensão. Estações elevatórias com capacidades e alturas manométricas diversificadas. Dessa seleção, obteve-se, para alguns componentes de sistema, maior quantidade de dados e para outros menos.

Cabe ressaltar que embora haja semelhança técnica, sempre há diferenças entre projetos. Há diferenças topográficas, geológicas, de pavimentos, de alguns materiais

ou equipamentos, sejam qualitativas ou quantitativas. É impossível conseguir homogeneidade total nos dados, seja com relação a tipos de serviços, seja com as quantidades deles em projetos de mesma capacidade. As condições locais, de natureza física ou econômica, podem ser diferentes. Embora tal diversidade provoque alguma variabilidade nos dados de custo, ela é importante para o objetivo da pesquisa, pois ela visa estimar custos de investimento a partir de parâmetros primários. Entretanto, alguns dados que apresentam custos fora da normalidade são avaliados e se necessário são eliminados. Como por exemplo, estações elevatórias de poço seco, estações elevatórias com fundações profundas, redes coletoras com volume excessivo de escavação em rocha, rede coletora de pequeno diâmetro com escoramento metálico-madeira em quantidade muito grande, que acarretariam distorções nos dados de custo. Considerou-se, para efeito desse estudo, os custos obtidos na época da efetiva realização do projeto.

Na pesquisa, estabelece-se uma tipologia de projetos de Sistemas de Esgoto Sanitário para identificação e avaliação dos dados de acordo com os seguintes aspectos:

- similaridade, pois tratam-se de projetos de uma mesma empresa, que procura manter padrões de construção, de materiais e de equipamentos e tem sua própria regulamentação para elaboração de seus custos de construção;
- funcionalidade e confiabilidade dos dados, pois os projetos de cada parte componente do sistema são executivos e completos;
- atualidade dos dados, os projetos consultados são aqueles elaborados nos últimos dez anos (do ano de 1992 a 2002) e que apresentam pouca mudança nos padrões, normas e tecnologias;
- homogeneidade física e econômica, dos projetos consultados cerca de 80% deles são de sistemas de esgoto de cidades litorâneas e 20% de cidades do interior, as quais apresentam diversidade de características topográficas, geológicas e econômicas.

Os projetos de engenharia da CASAN para implantação dos sistemas, tem seus custos orçados de acordo com o seu Manual de Regulamentação de Preços e Critérios de Medição. O mesmo manual citado no trabalho de JUNGLES (1994). A sua regulamentação tem por objetivo definir quais os serviços e materiais que estão

inclusos nos preços especificados em orçamentos das obras a empreitar. Ele visa também proporcionar o entendimento entre a fiscalização e o construtor no que se refere às medições de serviços para pagamento.

Todo o custo unitário ou global, referido no manual como preço unitário ou global de serviço, inclui em sua composição os custos relativos a materiais, mão-de-obra, veículos, equipamentos, ferramentas, aparelhos, instrumentos, água, esgoto, energia elétrica, segurança e ônus diretos e indiretos necessários à implantação do projeto.

Esses custos detalhados de serviços são apropriados para cada parte componente de sistema e discriminados por obras civis, materiais e equipamentos hidráulicos, conforme mostrado na tabela 3.1., a seguir.

São apresentados, portanto, nessa tabela, os custos diretos de construção com o BDI (Bonificações e Despesas Indiretas). Nela não estão incluídos, custos de projetos e taxas de administração, custos em desapropriações de terrenos e em compra de imóveis, pois estes dois últimos custos são ônus das prefeituras municipais onde são implantados os sistemas.

Tabela 3.1. Discriminação de custos de construção por componentes de sistema de esgoto sanitário.

Partes do Sistema	Obras Civis	Equipamentos Hidráulicos	Materiais Hidráulicos	Total
Rede Coletora				
Ligações Prediais				
Estação Elevatória				
Emissário				
Estação de Tratamento				

Os custos de projetos que fazem parte do trabalho, coletados na CASAN, são aqueles utilizados no ambiente da empresa, os quais são usados para pedidos de financiamento e de valor base para licitar as obras.

Dos projetos são selecionados dados para construção de modelos paramétricos de custo de construção para cada parte de sistema abaixo listada:

- rede coletora de esgoto;
- ligações prediais de esgoto;
- estação elevatória de esgoto;
- emissário de esgoto;
- estação de tratamento de esgoto.

Além dos dados de custo, são identificados e coletados outros dados. São características físicas, funcionais ou de capacidade dessas partes de sistema. Elas são definidas pela disponibilidade de suas informações no banco de dados formado e pela presença nas etapas iniciais de planejamento do empreendimento. Tais como: população beneficiada, densidade populacional e área urbana atendida pelo projeto, vazão de projeto, altura manométrica, extensão de tubulação, número de ligações prediais. Tais dados são os denominados direcionadores de custo e devem ser avaliados adiante, quando da modelação da relação paramétrica para cada componente de sistema.

3.2.3 (C) - Normalização dos Dados.

A normalização visa basicamente a homogeneidade de produção, de produtividade e econômica relativamente aos custos de implantação dos projetos.

Os dados históricos selecionados, como relatado anteriormente, tem origem nos projetos da CASAN, a qual tem mantido seus padrões tecnológicos de produtividade e produção com poucas alterações de processo. Já no aspecto econômico, referente às correções, devidas à inflação, os dados de custo merecem um trabalho de normalização.

A análise de custos de insumos e serviços, no Brasil, tem sido dificultada pelo processo inflacionário e pelas freqüentes alterações de padrão monetário. Entretanto, a partir de agosto de 1994, após um choque econômico com mudança de moeda, iniciou-se um período de estabilidade, com baixos índices inflacionários e com uma moeda forte em relação ao dólar. Isso trouxe facilidades na análise dos dados de custo da presente pesquisa.

Os dados de custos das partes de sistema de esgoto sanitário foram selecionados entre os diversos projetos elaborados, no período de 1992 até 2002. Apenas 25% deles são de projetos elaborados antes de agosto de 1994, portanto poucos projetos fazem parte daquele período conturbado da economia brasileira.

Para homogeneização dos valores monetários dos custos, é ideal o uso de um índice específico para atualização dos custos de saneamento. Na falta desse índice é utilizado o índice econômico da Fundação Getúlio Vargas, denominado IGP-DI (Índice Geral de Preço – Disponibilidade Interna). Portanto, todos os dados de custo são conduzidos para uma única data base (julho de 2000) e convertidos para o dólar americano equivalente, isto é, \$1,00 igual a R\$ 1,78. A escolha do dólar para conversão dos custos está baseada na sua estabilidade como moeda, pois apresenta baixíssima variabilidade relacionada aos problemas de inflação. Entretanto, alguns cuidados devem ser tomados, nas estimativas, quando o Real sofre grandes desvalorizações num curto espaço de tempo. Numa situação dessas, deve-se esperar até que o mercado absorva essa variação do Real em relação ao dólar, para evitar distorções nos resultados de custo estimados pelo modelo. Caso ela perdure, é importante que se use um índice diferente do dólar, que se identifique com as variações de custo do mercado, para atualizar a estimativa do modelo de julho de 2000 para data atual.

Com os dados de custo homogeneizados, o passo seguinte é o processo de identificação e análise dos direcionadores e o desenvolvimento das relações paramétricas de custo de cada componente de Sistema de Esgoto Sanitário.

3.2.4 (D) - Reconhecimento e Análise de Direcionadores de Custo.

Numa relação paramétrica, o custo estimado de um projeto é função de uma ou mais variáveis relevantes independentes. Estas variáveis são os direcionadores de custo, que são características ou propriedades técnicas de um projeto sendo estimado.

Rede Coletora de Esgoto - Para desenvolver relações paramétricas de custo para rede coletora, é importante o conhecimento de como ela é executada, quais custos incorrem na sua implantação e que dados geométricos, físicos, ou outros quaisquer de ordem primária possam se relacionar com seu custo de construção.

Ela é conhecida como uma rede de tubulações subterrâneas interligadas, implantadas, nas ruas, com profundidade mínima definida por norma, para escoar por gravidade os esgotos sanitários coletados das edificações de uma área urbana (cidade, vila, bairro, etc.).

O custo de execução da rede tem por base o processo construtivo convencional, que compreende: remoção da pavimentação da rua, abertura da vala, escoramento simples, esgotamento, assentamento da tubulação, construção de poços de visita, reaterro e reposição da pavimentação.

Com o intuito de dar homogeneidade aos custos de rede coletora, aqueles custos relativos aos serviços de escoramento metálico-madeira, ao rebaixamento do lençol freático, aos grandes volumes de escavação em rocha e as grandes extensões de pavimentação asfáltica, que distorcem excessivamente os custos médios, são expurgados do banco de dados de custos coletados.

Vários dados de projeto de rede relacionáveis com o seu custo são selecionados. Dentre eles, o diâmetro da tubulação, a extensão da rede, a área urbana de abrangência da rede, a população de saturação da área urbana e a densidade populacional da área. Todos são considerados direcionadores de custo, pois é fácil entender que cada um deles ou em conjunto explicam a variação do custo da rede. Quando o diâmetro da tubulação aumenta o custo da rede também aumenta, se sua extensão cresce o seu custo também cresce, e o mesmo acontece positivamente para os demais direcionadores em relação ao custo da rede. Alguns desses parâmetros são mais simples e de mais baixo custo de se obter, já outros, como o diâmetro das tubulações exige a elaboração detalhada do projeto. A área, a população e a densidade são dados básicos, obtidos já na fase de estudo da rede coletora. A extensão, embora pareça um dado básico, mesmo que possa ser associada ao comprimento de ruas onde a rede for implantada, apresenta discrepância em relação a esse comprimento. No caso da área de abrangência, basta conhecer a área urbana que se quer atender com rede de esgoto. Quanto à população de saturação e à densidade, apenas o estudo demográfico é suficiente.

Ligação Predial de Esgoto Sanitário - A ligação predial é uma tubulação que interliga o ramal predial de uma edificação à rede coletora de esgoto sanitário. Implantada nos diâmetros de 100mm e 150mm, ela ocorre em grande quantidade

no menor diâmetro(varia de 60 a 100%). Apenas nas áreas urbanas mais “verticalizadas” (onde há mais edifícios com mais de oito apartamentos) é que há um aumento significativo de ligações em 150mm.

O seu comprimento médio padronizado no projeto é de 8,0m. Cerca de 99% dos projetos de ligações usam material em PVC.

O custo da ligação compreende efetivamente: remoção e reposição de pavimentação, escavação e reaterro, aquisição e assentamento de tubos e conexões e caixa de inspeção com aquisição do tampão em ferro fundido.

Na coleta dos dados históricos, foram selecionados e analisados vários direcionadores de custo, tais como: população inicial beneficiada, população final beneficiada, área urbana beneficiada, número de ligações em 100mm, número de ligações em 150mm e numero total de ligações. Dentre essas variáveis, a que atende o objetivo da pesquisa e também tem melhor identidade com os custos é o número total de ligações em cada projeto. Embora as variáveis numero de ligações em 100 e 150mm apresentem igualmente bom relacionamento com os custos, não são selecionadas para o modelo, pois não há como definir a proporção delas na fase de planejamento. Além disso, observando os projetos, as áreas urbanas consideradas apresentam, em média, baixa concentração de edifícios com mais de 8 apartamentos

Estação Elevatória de Esgoto Sanitário - Estação elevatória é uma unidade de sistema de esgoto sanitário, basicamente constituída de uma estrutura de concreto enterrada, denominada poço, de onde os conjuntos motobombas recalcam o esgoto de um nível para outro.

Seus custos expressam os gastos com os insumos, mão-de-obra, materiais e equipamentos, necessários à implantação de uma unidade de estação elevatória, cujos conjuntos motobombas são em sua maioria submersíveis.

Salvo pequenas alterações, essas estações são padronizadas. Elas compreendem um poço circular em concreto com profundidade de 4,00 a 8,00m onde ficam alojados os conjuntos motobombas; um poço de gradeamento; as tubulações de sucção e do barrilete; a estrutura de içamento das motobombas e a estrutura do quadro de comando.

Os referidos custos incluem, escavação, escoramento, esgotamento e rebaixamento do lençol freático, reaterro, concreto em geral, revestimentos, impermeabilização, instalações hidráulicas, aquisição e montagem de equipamentos elétricos e eletromecânicos, tubulações e testes de partida.

Vários dados de projeto de estação elevatória apresentam algum relacionamento com seu custo. Dentre eles, são selecionados, a potência consumida, a capacidade de vazão de projeto da estação e a altura manométrica de recalque. Todos são considerados direcionadores de custo, pois é simples de entender que cada um deles ou em conjunto explicam a variação de custo da elevatória. É possível observar que quando a potência aumenta há o aumento do custo da estação. O mesmo acontece quando cresce a capacidade ou altura manométrica, o custo também cresce.

Emissário de Esgoto Sanitário - É definido como a tubulação utilizada para transportar o esgoto sob pressão, seja por gravidade ou através do recalque por bombeamento.

Seu custo é caracterizado pelo seguintes serviços: locação e cadastro, limpeza do terreno, remoção e reposição de pavimentação, escavação, reaterro, substituição de solos, escoramentos, esgotamento e rebaixamento do lençol freático, embasamentos, aquisição e assentamento de tubos, peças e conexões e testes de estanqueidade.

Vários dados de projeto de emissário relacionáveis com o seu custo foram selecionados. Dentre eles, o diâmetro da tubulação, a extensão do emissário e a vazão de projeto. Todos são considerados direcionadores de custo, pois é fácil entender que cada um deles ou em conjunto explicam a variação do custo do emissário. Quando o diâmetro da tubulação aumenta o custo do emissário também aumenta, se sua extensão cresce o seu custo também cresce, e o mesmo acontece positivamente para a vazão de projeto e o custo do emissário. Alguns desses parâmetros são mais simples e de mais baixo custo de se obter, já outros, como o diâmetro da tubulação exige a elaboração detalhada do projeto. A vazão de projeto e a extensão são dados básicos, obtidos já na fase de estudo do emissário.

Estação de Tratamento de Esgoto Sanitário - É a parte de sistema de esgoto sanitário que recebe todo esgoto coletado promovendo a remoção da matéria

orgânica desse efluente, em níveis que variam de 50 a praticamente 100%, antes que ele seja encaminhado a um corpo receptor (rio, lago ou mar). Pode ser construída em terra, concreto ou aço.

Há vários tipos de estações de tratamento, dentre os disponíveis na CASAN são coletados dados de apenas quatro projetos de estações de lodos ativados em condições de sistematização e normalização.

O custo da estação é caracterizado, no caso de lodos ativados, pelos seguintes serviços: locação, limpeza do terreno, escavação, aterro, escoramentos, esgotamento, execução de estruturas de concreto, acabamentos, impermeabilização, aquisição, montagem de equipamentos, de tubos, peças e conexões e testes de funcionamento.

Vários dados de projeto de estação de tratamento relacionáveis com o seu custo podem ser selecionados. Entretanto, os mais fáceis e simples de obter, já nas primeiras fases, são a população geradora dos esgotos e a vazão de projeto. Essas variáveis são consideradas direcionadores de custo. Observe-se que quando eles aumentam o custo da estação também aumenta.

3.2.5 (E) - Desenvolvimento das Relações do Modelo Paramétrico de Custo.

Rede Coletora de Esgoto Sanitário – São elaboradas análises expeditas de estatística de ajuste com o intuito de escolher as variáveis que melhor direcionem os custos de rede para o seu modelo de relação paramétrica,. Todas os direcionadores, também ditos variáveis independentes, são relacionados logaritmicamente, separados e em conjunto, com o custo de rede coletora, conforme ilustra a tabela adiante. A área de abrangência (A) se relaciona através do $\log(A)$, a densidade populacional pelo $\log(D)$ e a população de saturação pelo $\log(P)$.

Tabela 3.2. Análise comparativa entre os possíveis ajustes do custo de rede coletora e seus direcionadores.

Erro quadrado médio	R^2	R^2	Cp	Variáveis
0,143803	96,89	96,48	2,70	$\log(D).\log(P)$
0,143829	96,89	96,48	2,70	$\log(A).\log(P)$
0,143852	96,89	96,48	2,71	$\log(A).\log(D)$
0,146668	97,04	96,41	4,0	$\log(A).\log(D).\log(P)$
0,299394	93,10	92,67	18,66	$\log(P)$
0,308931	92,88	92,44	19,70	$\log(A)$

A variável ou a combinação de variáveis independentes que melhor explica o custo baseia-se no menor valor do erro quadrado médio, no menor valor Cp de Mallows e no mais alto valor de R^2 . Logo, ao se observar a tabela acima, ela indica que as três primeiras combinações de variáveis são as melhores. No entanto, desenvolvendo os testes de significância é possível observar que a melhor combinação envolve a área de abrangência da rede coletora e densidade populacional da área.

O modelo da expressão de custo, estudada para a rede coletora, está fundamentado nas duas variáveis independentes e explanadoras citadas anteriormente: a área de abrangência da rede coletora, medida em hectare (ha) e a densidade populacional de saturação da área atendida, medida em habitantes por hectare (hab/ha).

Ligação Predial de Esgoto Sanitário - Na tabela 3.3, onde $\log(P)$ é o logaritmo da população beneficiada inicial, $\log(PF)$ é o da população beneficiada final, $\log(A)$ é o da área urbana beneficiada e $\log(L)$ é o do número total de ligações, é possível afirmar que o número de ligações é a variável independente que melhor se correlaciona com os custos.

Tabela 3.3. Análise comparativa entre os possíveis ajustes do custo de ligações prediais e seus direcionadores.

Erro quadrado médio	R^2	R^2	Cp	Variáveis
0,135071	94,71	94,18	0,028	$\log(L)$
0,144562	94,91	93,77	1,73	$\log(L).\log(P)$
0,145591	94,87	93,73	1,78	$\log(L).\log(A)$
0,149672	94,73	93,56	2,00	$\log(L).\log(PF)$
0,150387	95,29	93,52	3,15	$\log(L).\log(P).\log(PF)$
0,161982	94,93	93,03	3,70	$\log(L).\log(P).\log(A)$

Estação Elevatória de Esgoto Sanitário - Para escolher as variáveis melhor direcionadoras dos custos de estação elevatória para o seu modelo de relação paramétrica, são elaboradas análises expeditas de estatística de ajuste. Todos os direcionadores, também ditos variáveis independentes, são relacionados logaritmicamente, separados e em conjunto, com o custo de estação elevatória, conforme ilustra a tabela abaixo. A potência consumida da estação elevatória (P) se relacionou através do $\log(P)$ com o $\log(C)$ do custo, a capacidade por $\log(Q)$ e a altura manométrica por $\log(H)$.

Tabela 3.4. Análise comparativa entre os possíveis ajustes do custo de estação elevatória e seus direcionadores.

Erro quadrado médio	R^2	R^2	Cp	Variáveis
0,101607	94,38	94,01	2,22	$\log(Q).\log(H)$
0,106636	94,10	93,71	3,67	$\log(P).\log(Q)$
0,104284	94,42	93,85	4,00	$\log(P).\log(Q).\log(H)$
0,136626	92,19	91,94	11,61	$\log(Q)$
0,14646	91,90	91,36	15,13	$\log(P).\log(H)$
0,174821	90,01	89,69	22,96	$\log(P)$

A variável ou a combinação de variáveis independentes, que melhor explica o custo, baseia-se no menor valor do erro quadrado médio, no menor valor Cp de Mallows e no mais alto valor de R^2 . Logo, na tabela acima, se observada, ela indica que as três primeiras combinações de variáveis são as melhores. No entanto, ao desenvolver os testes de significância, é possível observar que a melhor relação com o custo envolve a combinação da capacidade de recalque (Q) com a altura manométrica (H).

Emissário de Esgoto Sanitário - São elaboradas análises expeditas de estatística de ajuste para verificar-se as melhores relações com seus custos. Todos os direcionadores, também ditos variáveis independentes, são relacionados logaritmicamente, separados e em conjunto, com o custo de emissário, conforme é apresentado abaixo. A vazão de projeto (Q) se relaciona através do $\log(Q)$, a extensão (E) por $\log(E)$ e o diâmetro (D) por $\log(D)$.

Tabela 3.5. Análise comparativa entre os possíveis ajustes do custo de emissário e seus direcionadores.

Erro quadrado médio	R^2	R^2	Cp	Variáveis
0,182328	96,03	95,75	2,80	$\log(Q).\log(E)$
0,183631	96,15	95,72	4,00	$\log(Q).\log(D).\log(E)$
0,253802	94,48	94,09	13,70	$\log(D).\log(E)$
1,00298	78,20	76,65	127,93	$\log(Q).\log(D)$

A variável ou a combinação de variáveis independentes que melhor explica o custo baseia-se no menor valor do erro quadrado médio, no menor valor Cp de Mallows e no mais alto valor de R^2 . Logo, se observada a tabela acima, ela indica que as três primeiras combinações de variáveis são as melhores. Porém, ao se desenvolver os testes de significância é possível observar que a melhor relação com o custo é a combinação que envolve a vazão de projeto e a extensão de emissário.

Estação de Tratamento de Esgoto Sanitário - São elaboradas análises expeditas de estatística de ajuste para escolher as variáveis que melhor direcionam os custos de estação de tratamento. Todos os direcionadores são relacionados separados e em conjunto com o custo de estação de tratamento, conforme é apresentado abaixo. A vazão de projeto se relaciona através de Q e a população por P.

Tabela 3.6. Análise comparativa entre os possíveis ajustes do custo de estação de tratamento e seus direcionadores.

Erro quadrado médio	R^2	R^2	Cp	Variáveis
$1,3 \times 10^{10}$	99,35	98,04	3,0	$(Q).(P)$
$2,2 \times 10^{11}$	77,80	66,70	34,0	(Q)
$3,7 \times 10^{11}$	62,92	44,38	56,9	(P)

A variável ou a combinação de variáveis independentes que melhor explica o custo, baseia-se no menor valor do erro quadrado médio, no menor valor C_p de Mallows e no mais alto valor de R^2 . Se observada a tabela 3.6, ela mostra que apenas a primeira relação de variáveis é a melhor. Mas, ao se desenvolver os testes de significância é possível observar se a melhor relação com o custo é significativa no modelo.

3.2.6 (F) – Modelação Matemática.

Uma vez o banco de dados desenvolvido e uma hipótese definida, a relação paramétrica de custo está pronta para ser modelada matematicamente.

A relação de estimativa pode variar em complexidade, a partir algo bastante simples; como uma expressão numérica de valor, até algo mais complexo; como uma expressão matemática multivariável.

A modelação pode passar de um simples fator ou razão a análises de regressão, do tipo linear simples, linear múltipla ou não linear.

A regressão linear é definida como uma busca da melhor relação entre duas variáveis, através do estabelecimento de uma equação matemática linear (reta).

Caso se queira prever os valores de uma variável que esteja relacionada com outra, cujos valores sejam previamente conhecidos, deve ser usada uma equação de regressão.

A regressão linear significa que a relação funcional das variáveis pode ser descrita graficamente (sobre um sistema de coordenadas X-Y) por uma linha reta e matematicamente através da equação:

$$Y = A + BX, \text{ onde :}$$

Y representa o valor estimado da variável dependente,

X é a variável independente,

B é a declividade da reta ou também coeficiente angular,

A é o ponto ou cota onde a reta intercepta o eixo vertical (Y).

A equação de regressão de duas variáveis consiste de duas partes, a parte funcional e a parte aleatória. Assim, uma população pode ser representada pela equação $Y = A + BX + E$. Onde, a porção " $A + BX$ " é a parte funcional (uma linha

reta), e E (o termo do desvio) é a parte aleatória. A parte do resíduo (E) está sempre presente por causa dos erros de determinação, medida e observação. Esses tipos de erros existem com freqüência devido às limitações humanas e às limitações associadas aos eventos do mundo real (DEPARTMENT OF DEFENSE,1999).

Visto que é praticamente impossível obter os dados de uma população inteira, normalmente trabalha-se com uma amostra dessa população. Assim, trabalha-se com uma amostra para adaptar a equação de regressão para a forma : $Y = a + bX + e$. Portanto, novamente o termo " $a + bX$ " representa a parte da função da equação e " e " a parte aleatória. A estimativa dos parâmetros verdadeiros " A " e " B " da população é representada na equação da amostra por " a " e " b " respectivamente. Dessa forma, " a " e " b " são estatísticos. Isto é, eles são estimativas dos parâmetros da população. Como parâmetros estatísticos, eles estão sujeitos aos erros de amostragem. Conseqüentemente, um bom planejamento de amostras aleatórias é importante (DEPARTMENT OF DEFENSE,1999).

A aproximação matemática denominada método dos mínimos quadrados é a raiz da análise de regressão, que pode ser definida como a natureza matemática da associação entre duas variáveis. Essa técnica, usada para melhorar a habilidade para prever o valor mais próximo do real da variável dependente, especifica uma linha que melhor ajusta o conjunto de dados. O método faz isso minimizando a soma dos quadrados dos desvios que resultam da diferença entre os valores observados de Y e os valores calculados de Y . O valor observado representa o valor realmente existente, na base de dados, e o valor calculado, identificado como Y_c , é aquele valor estimado pela equação tomando o mesmo valor de X . A diferença entre os valores observados e calculados é representada pelo termo " e " (desvio) da equação linear (DEPARTMENT OF DEFENSE,1999).

A técnica dos mínimos quadrados analisa cada par (X, Y) na base de dados, refinando os parâmetros para os termos da declividade e do intercepto vertical, até encontrar uma equação para a linha que minimiza a soma dos quadrados dos desvios.

Para calcular a linha de melhor ajuste dos mínimos quadrados para uma base de dados de n números de observações é necessário encontrar a função $(a + bX)$ que minimize:

$$\min \sum_{i=1}^n e_i^2 = (Y_1 - Y_{C1})^2 + (Y_2 - Y_{C2})^2 + (Y_3 - Y_{C3})^2 + \dots + (Y_n - Y_{Cn})^2$$

Os valores de a e b para a reta $Y_c = a + bX$, que minimiza a soma dos quadrados dos desvios, são obtidos da solução de um sistema de equações normais (STEVENSON, 1981):

$$(1) SY = na + b(SX)$$

$$(2) SXY = a(SX) + b(SX^2)$$

Referências contidas em muitos livros de estatística, mostram que essas duas equações satisfazem as exigências da regressão comum dos mínimos quadrados.

As propriedades da técnica são:

- ela considera todos os pontos de dados;
- a soma dos quadrados dos desvios entre a linha e os pares de dados observados é o valor mínimo possível.

Similaridades entre estas duas propriedades e a média aritmética são observadas. A média aritmética é a soma dos valores da variável independente divididos pelo número de observações e a soma dos valores da variável dependente dividida pelo número de observações. Isso resulta que o ponto (\bar{Y}, \bar{X}) relativo às médias cai sobre a reta dos mínimos quadrados.

Das equações normais são derivadas duas equações para calcular a e b diretamente :

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$a = \bar{Y} - b\bar{X}.$$

(Lembrando que uma vez conhecido b , é possível resolver a equação da reta. $Y = a + bX$ para a , porque é sabido que o ponto (\bar{Y}, \bar{X}) deve cair sobre a reta.)

(DEPARTMENT OF DEFENSE, 1999).

Então, obtidos os valores de a e b , a equação da reta que melhor ajusta uma base de dados está definida.

Note-se que a regressão linear não testa se seus dados são lineares. Ela assume que seus dados são lineares e busca os coeficientes angular b e o linear a que produzam uma linha reta que melhor ajuste seus dados

Logo, segundo (MOTULSKY, 1999), quando trabalha-se com a análise de regressão linear tem-se a considerar um certo número de limitações, erros e cuidados, que a seguir serão analisados.

- Em muitas pesquisas, a relação entre X e Y é uma linha curva e, portanto, produz a regressão linear inadequada. Ou os dados podem ser transformados ou então usa-se o ajuste da regressão não linear.
- A análise de regressão linear assume que a dispersão dos pares de dados ao redor da linha é Gaussiana.
- A variância de cada conjunto da variável independente é constante, portanto, a hipótese é violada se os pontos com valores altos ou baixos de X tenderem a ficar mais distante da linha de melhor ajuste.
- Se um ponto está acima ou abaixo da linha é por mero acaso, e não influencia se outro ponto está em quaisquer dessas posições. Quando a hipótese é violada, os dados são autocorrelacionados. Essa hipótese requer que o termo do desvio seja uma variável verdadeiramente aleatória.
- O valor médio de cada distribuição cai sobre a linha de regressão.
- O modelo de regressão especifica que a variável independente deve ser um número fixo e não uma variável aleatória.
- Uma verificação lógica deve ser feita, determinando uma hipótese apropriada e analisando a base de dados de forma que uma avaliação possa ser relacionada a causa e efeito.
- Se a população base não é mais relevante devido as mudanças, na tecnologia, por exemplo, então a equação de regressão linear pode não ser a melhor ferramenta de previsão para usar. Ao se usar um modelo de regressão

deve-se garantir que os dados não se tornaram obsoletos para previsão no momento atual ou no futuro.

Já a regressão múltipla é basicamente a associação de três ou mais variáveis e nada mais que uma extensão da regressão linear simples. Continua havendo apenas uma variável dependente, entretanto, passa-se a ter duas ou mais variáveis independentes (ditas também explanatórias).

A análise de regressão múltipla assume que as variações em y podem ser melhor explicadas associando-a a mais de uma variável independente.

A finalidade das variáveis independentes adicionais é melhorar a capacidade de predição se comparada com a regressão linear simples (STEVENSON, 1981).

A equação de regressão múltipla pode exprimir y em função de variáveis independentes x_1, x_2, \dots e ser denominada equação de regressão de y para x_1, x_2, \dots

Por exemplo, no caso de três variáveis independentes a equação tem a seguinte forma:

$$Y_c = a + b_1x_1 + b_2x_2 + b_3x_3,$$

em que a, b_1, b_2, b_3 são constantes, sendo:

y_c = o valor estimado da variável dependente;

a = o intercepto vertical, quando $x_1, x_2, x_3 = 0$;

x_1 = a primeira variável explanatória;

b_1 = o coeficiente angular da linha relativo às variações em x_1 ;

x_2 = a segunda variável explanatória;

b_2 = o coeficiente angular da linha relativo às variações em x_2 ;

x_3 = a terceira variável explanatória;

b_3 = o coeficiente angular da linha relativo às variações em x_3 (DEPARTMENT OF DEFENSE, 1999) (STEVENSON, 1981).

Encontrar as combinações das variáveis explanatórias não é tarefa fácil. Em alguns casos, resultados de regressão múltipla podem parecer paradoxais. O modelo ajusta bem os dados, embora nenhuma das variáveis independentes tenham estatisticamente um impacto significativo na predição de y . Quando duas variáveis independentes apresentam alta correlação, ambas carregam essencialmente a mesma informação. Nesse caso, nenhuma pode contribuir significativamente para o modelo depois que a outra é incluída, mas já quando se

apresentam juntas elas contribuem muito. Removendo do modelo, as duas variáveis, o ajuste seria muito pior. Assim, o modelo global ajusta bem os dados, mas nenhuma variável x faz uma contribuição significativa quando é adicionada ao seu modelo final. Quando isso acontece as variáveis independentes são colineares e os resultados mostram multicolinearidade (DEPARTMENT OF DEFENSE, 1999).

Se o objetivo é simplesmente prever y de um conjunto de variáveis x , então a multicolinearidade não causa problema. As previsões ainda serão precisas.

Usando o Software Statgraphics, que faz uso de métodos estatísticos, os modelos são desenvolvidos e encontram melhor ajuste através da técnica de regressão linear múltipla, onde são obtidas transformadas logarítmicas das equações matemáticas representativas.

Basicamente os modelos das partes componentes de sistema de esgoto sanitário são expressões do tipo:

$$Y = K X_1^a X_2^b$$

Onde:

Y = custo do componente de sistema de esgoto sanitário (\$ USA);

K = constante de custo;

X_1 = a primeira variável independente (direcionador de custo);

a = expoente de X_1 que representa o seu fator de economia de escala;

X_2 = a segunda variável independente (direcionador de custo);

b = expoente de X_2 que representa o seu fator de economia de escala.

Para o ajuste do modelo, utilizando a técnica de regressão linear múltipla, é necessário trabalhar com a transformada logarítmica da equação acima, representada da seguinte forma:

$$\log(Y) = \log(K) + a.\log(X_1) + b.\log(X_2)$$

3.2.7 (G) – Testes de Significância do Modelo Ajustado.

Depois de discutir a técnica de regressão, volta-se para avaliar a qualidade da estimativa paramétrica de custos. Assim, deve-se responder as perguntas, conforme DEPARTMENT OF DEFENSE (1999):

- Uma equação de estimativa paramétrica de custo é significativamente boa e está significativamente boa para estimar o custo de itens específicos ou serviços?
- Qual é o intervalo de confiança da estimativa (i.e., como provavelmente é o custo estimado para se cair dentro de uma faixa específica de resultados de custo)?

Para avaliar a qualidade de uma relação paramétrica de custo, há uma relação de itens para examinar. Há elementos que focam na validação estatística da relação paramétrica ou modelo, tais como: a qualidade dos dados, a relação lógica entre eles, os testes estatísticos t e F, o desvio padrão, o coeficiente de variância e o coeficiente de determinação (R^2). Enquanto outros, como o número de observações, grau de liberdade, “outliers” e amplitude dos dados, enfocam no uso da relação ou modelo para predizer estimativas futuras.

O exame estatístico começa tipicamente com uma avaliação das variáveis individuais no modelo. O método mais comum para avaliar a significância da variável na relação é o teste t para cada variável explanatória. O próximo passo a avaliar é a significância da equação inteira. O teste F é a estatística mais comum para avaliar esta qualidade. Assumindo que as variáveis e a equação inteira têm significância, o passo seguinte é julgar o tamanho e a proporção do desvio estimado da equação. O desvio padrão da estimativa (DPE ou DP) e o coeficiente de variância (CV) fornecem esta compreensão.

Finalmente, a análise estatística típica conclui com o exame do valor do coeficiente de determinação (R^2) ou R^2 ajustado ao comparar os modelos com um número diferente de variáveis independentes para cada modelo. O coeficiente de determinação mede a porcentagem de variação da variável dependente explanada pelas variáveis independentes.

3.2.8 (H) – Validação do Modelo.

Visando determinar se as relações paramétricas construídas são boas para predição dos custos de partes componentes de sistema de esgoto sanitário, é necessário avaliar a precisão delas. Portanto, são separados dados independentes do conjunto de dados históricos que geraram os modelos paramétricos. Para cada

parte de sistema, são coletados tantos dados quanto disponíveis, desde que o conjunto de dados formadores de cada modelo não sejam considerados insuficientes para sua construção.

De posse dos dados independentes de projetos existentes, são estimados os custos através dos modelos e comparados com os custos desses projetos. Para efeito de comparação, os dados de custo são saneados dos mesmos efeitos que pode distorcer o padrão estabelecido no processo de seleção dos dados históricos, e normalizados para a mesma data dos modelos, isto é, para julho de 2000 e convertidos para o dólar americano, a mesma moeda utilizada pelos modelos.

Os resultados dos modelos mostram, quando comparados com os custos dos projetos, o nível de precisão de cada um deles.

4 RESULTADOS

Cerca de 50 projetos são consultados na CASAN, e deles são selecionados aqueles que apresentavam-se completos e executáveis, de maiores semelhanças técnicas e de diversos tamanhos e capacidades. Para essa avaliação, procurou-se projetos de pequenas redes coletoras com apenas o diâmetro mínimo e outras maiores que atingissem vários diâmetros de tubulações e vários quilômetros de extensão. Estações elevatórias com capacidades e alturas manométricas diversificadas. Dessa seleção, obteve-se, para alguns componentes de sistema, maior quantidade de dados e para outros menos.

Cabe ressaltar que, embora haja semelhança técnica, sempre há diferenças entre projetos. É impossível conseguir homogeneidade total nos dados, seja com relação a tipos de serviços seja com as quantidades deles em projetos de mesma capacidade. As condições locais, de natureza física ou econômica, podem ser diferentes. Entretanto, alguns dados que apresentam custos fora da normalidade são avaliados e se necessário são eliminados. Como por exemplo, estações elevatórias de poço seco, estações elevatórias com fundações profundas, redes coletoras com volume excessivo de escavação em rocha, rede coletora de pequeno diâmetro com escoramento metálico-madeira em quantidade muito grande, que acarretariam distorções nos dados de custo. Considerou-se, para efeito desse estudo, os custos obtidos na época da efetiva realização do projeto.

4.1 Modelo de Custo de Rede Coletora de Esgoto Sanitário

Na tabela a seguir estão apresentados os dados da amostra selecionada e normalizada para o ajuste do modelo de rede coletora. Estão apresentadas duas colunas de custos orçados, uma expressa em reais equivalentes ao mês de julho de 2000 [R\$(07/00)] e a outra em dólar americano (\$ USA). A função é ajustada pelo custo orçado em dólar. Uma terceira coluna, representa os custos por habitante.

Tabela 4.1. Dados de projetos de redes coletoras.

Número De Ordem	Área (ha)	Densidade (hab/ha)	Custo R\$ (07/00)	Custo (\$ USA)	Custo (\$ USA/hab)
1	4,44	45,05	30.084,58	16.873,01	84,36
2	3,57	105,00	32.883,33	18.442,70	49,20
3	4,52	155,75	57.545,91	32.274,77	45,85
4	4,06	230,00	101.790,04	57.089,20	61,14
5	14,59	137,08	295.512,79	165.739,09	82,87
6	74,80	66,84	764.256,52	428.635,18	85,73
7	72,31	100,00	844.614,70	473.704,26	65,51
8	47,72	362,34	1.392.652,66	781.072,72	45,17
9	88,54	249,34	1.516.679,20	850.633,31	38,53
10	263,00	41,86	1.687.689,24	946.544,72	85,98
11	148,59	24,23	1.802.874,93	1.011.146,90	280,85
12	492,00	112,12	3.335.495,27	1.870.720,85	33,91
13	146,60	331,49	4.636.366,94	2.600.317,97	53,51
14	186,12	574,23	4.695.586,04	2.633.531,15	24,64
15	307,40	287,79	5.800.373,73	3.253.154,08	36,77
16	342,84	290,18	7.252.609,34	4.067.644,05	40,89
17	587,57	148,66	13.099.366,00	7.346.812,11	84,11
18	665,75	184,59	23.897.241,36	13.402.827,46	109,06

Fonte: Projetos arquivados na Companhia Catarinense de Águas e Saneamento (CASAN)

Com o objetivo de definir o modelo de custo, é verificada a tendência inicial do comportamento da amostra. A variável dependente (custo) e as variáveis independentes da amostra, previamente selecionadas e associadas como direcionadoras dos respectivos custos, são plotadas em gráficos, em valores logarítmicos, nas figuras 4.1, 4.2, e 4.3 a seguir. Fazendo uma análise inicial dos gráficos, percebe-se as tendências dessas variáveis.

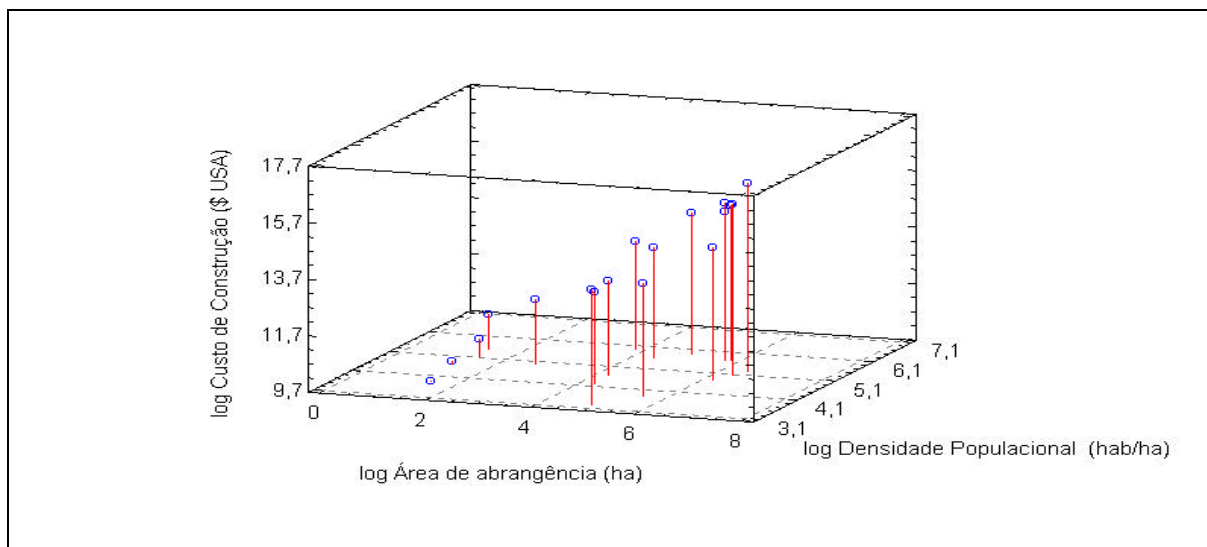


Figura 4.1. Custos de construção de redes coletoras em função de área de abrangência (ha) e densidade populacional (hab/ha).

Observa-se pelos pontos plotados na figura acima, com o crescimento da área de abrangência, que os custos de construção da rede também tendem a crescer. De igual modo, quando a densidade populacional aumenta, o custo da rede também tende a aumentar. Essas tendências mostram que as variáveis escolhidas para explicar os custos de rede coletora apresentam evidências de correlação com os custos de construção de modo individual e conjunto.

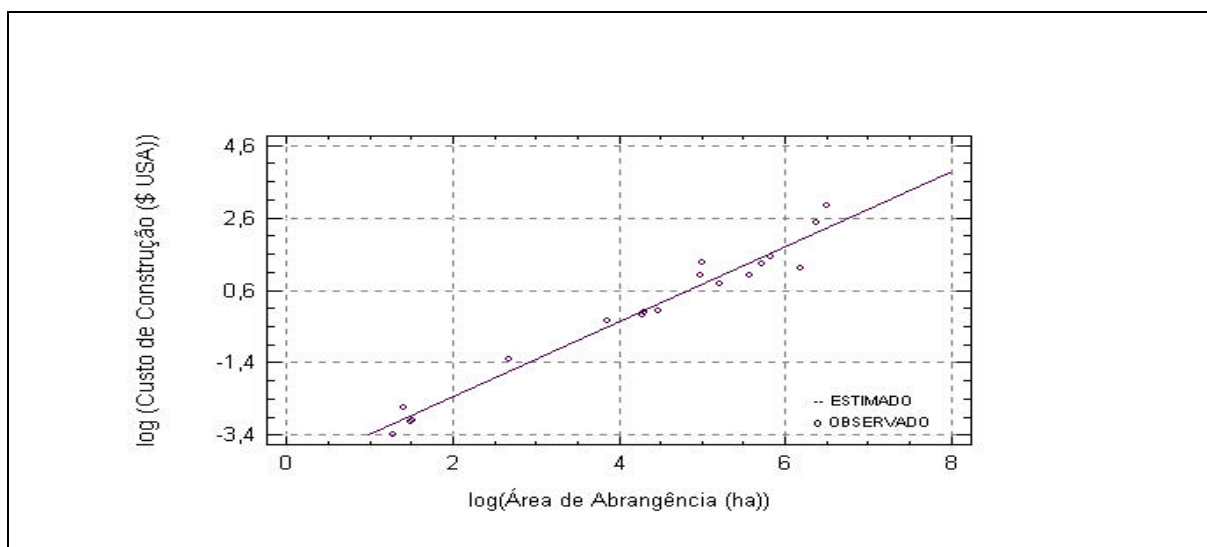


Figura 4.2. Custos de rede coletora em função de área de abrangência.

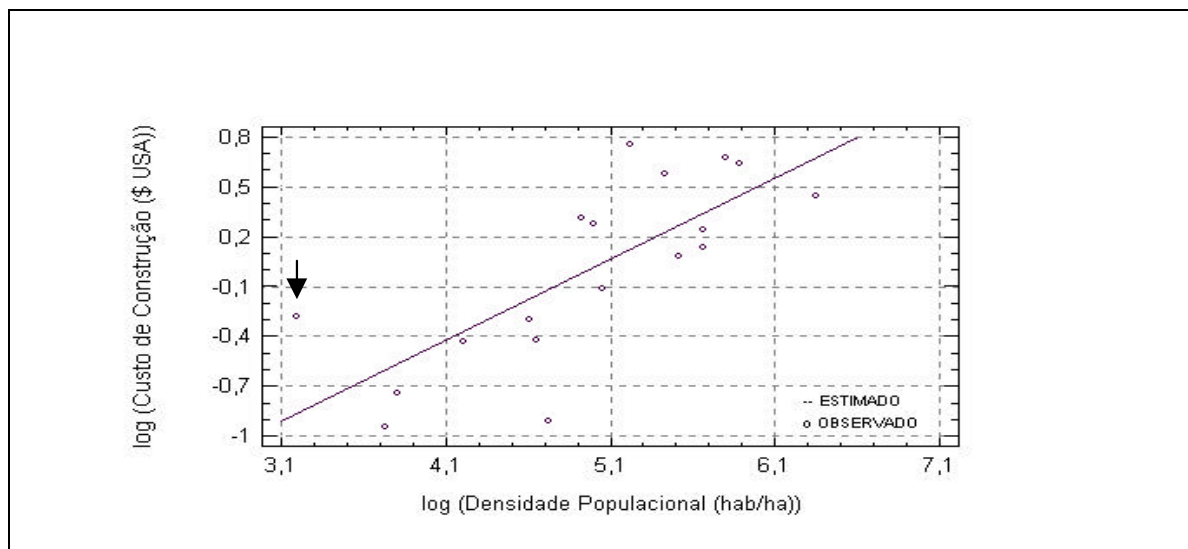


Figura 4.3. Custos de redes coletoras em função de densidade populacional (hab/ha).

Depreende-se em relação às correlações, observados os pontos plotados nos dois gráficos anteriores, que existe tendência de crescimento log-linear, podendo ter como resultante o traçado de uma reta. Entretanto, como trata-se de uma escala logarítmica, o modelo de custo em função das variáveis área e densidade é não linear.

A equação matemática adotada para expressar este modelo segue apresentada abaixo:

$$Y_1 = K_1 L_1^b X_1^a \quad (01)$$

Onde:

Y_1 = Custo de Rede Coletora de Esgoto Sanitário (\$ USA);

K_1 = Constante de Custo para Rede Coletoras;

L_1 = Densidade Populacional de Saturação na Área de Abrangência (hab/ha);

b = Expoente de L_1 que representa o Fator de Economia de Escala da Densidade Populacional da Área ;

X_1 = Área de Abrangência da Rede Coletora (ha);

a = Expoente de X_1 que representa o Fator de Economia de Escala da Área de Abrangência da Rede Coletora.

Utilizando o software Statgraphics faz-se a análise de regressão dos dados, bem como as várias análises estatísticas necessárias para avaliar os testes de

significância do modelo de custo ajustado. O nível de confiança escolhido para as análises estatísticas foi de 95%.

O exame estatístico começa tipicamente com uma avaliação das variáveis individuais no modelo. O método mais comum para avaliar a significância da variável na relação é o teste t para cada variável explanatória. O próximo passo a avaliar é a significância da equação inteira. O teste F é a estatística mais comum para avaliar essa qualidade. Assumindo que as variáveis e a equação inteira têm significância, o passo seguinte é julgar o tamanho e a proporção do desvio estimado da equação. O desvio padrão da estimativa (DPE ou DP) e o coeficiente de variância (CV) fornecem essa compreensão.

Finalmente, a análise estatística típica é concluída com o exame do valor do coeficiente de determinação (R^2) ou (R^2 ajustado) ao comparar os modelos com um número diferente de variáveis independentes para cada modelo.

Para desenvolver a análise de regressão linear múltipla do modelo é usada a transformada da equação (1):

$$\log(Y_1) = \log(K_1) + b.\log(L_1) + a.\log(X_1) \quad (2)$$

Onde:

$\log(Y_1)$ = logaritmo natural do custo de rede coletora de esgoto sanitário (\$ USA);

$\log(K_1)$ = logaritmo natural da constante de custo para rede coletora;

$\log(L_1)$ = logaritmo natural da densidade populacional de saturação na área de abrangência (hab/ha);

$\log(X_1)$ = logaritmo natural da área de abrangência da rede coletora (ha);

a e b = são os fatores de economia de escala já definidos na equação (1).

A tabela a seguir apresenta o resultado para o ajuste do modelo, isto é, a determinação da constante K_1 e dos fatores exponenciais de economia de escala (a e b) das variáveis independentes do modelo de custo para rede coletora.

Tabela 4.2. Ajuste do modelo de custo de rede coletora.

Parâmetros	Estimativa	Erro Padrão	t estatístico	Nível de Significância
$\log(K_1)$	6,5066	0,56959	11,4232	0,0000
a	1,0358	0,05088	20,3566	0,0000
b	0,4874	0,11077	4,4001	0,0005

Os valores dos testes mostram que os fatores de economia de escala (expoentes) são significativamente diferentes de zero. Isto é, ambas as variáveis independentes, área e densidade, são relevantes na avaliação de custo de rede coletora e se correlacionam positivamente com a variável dependente.

Os testes de significância da hipótese nula dessas variáveis analisados pela distribuição de Student, estão representados nas duas últimas colunas da tabela acima com níveis de significância praticamente nulos. Isso indica que a probabilidade dos fatores de escala das variáveis, definidos pela correlação, serem zero é igualmente nula. Se o expoente de uma variável é zero, qualquer valor que ela assuma terá como resultado sempre o número um. Assim, significa que a variável não influenciará os valores de custo resultantes do modelo.

Os resultados do erro padrão, o qual mede a precisão dos fatores de escala estimados em termos absolutos, mostram uma variação maior para o fator de escala da variável densidade populacional. Esse resultado demonstra menor precisão na determinação desse fator. Contudo, em relação ao fator de escala da variável área de abrangência o erro padrão é menor e representa uma variação de apenas 4,9% em torno da média.

Para a variância, estão representados os resultados da análise completa na tabela a seguir.

Tabela 4.3. Análise da variância do ajuste do modelo.

Fonte	Somatório dos Quadrados	Graus de Liberdade	Quadrado Médio	Razão F	Valor p
Modelo	67,3422	2	33,6711	234,07	0,0000
Resíduos	2,1578	15	0,1438		
Total	69,5000	17			

O teste de hipótese pela distribuição F de Fisher realizado para o modelo de custo, onde a Razão F é significativamente maior do que um, levou a rejeição da hipótese nula para um nível de significância de 5%. Isso representa que o modelo tem boa aderência à tendência dos dados observados.

Esse teste baseia-se na relação entre o quadrado médio devido à regressão e o quadrado médio residual para explicar a aceitação ou rejeição da hipótese nula. Quanto maior o F observado maior é a probabilidade de rejeição da hipótese nula.

Outros dados estatísticos para o modelo são apresentados abaixo:

Número de dados analisados	: 18
Número de dados desconsiderados	: 0
Coeficiente de determinação (R^2)	: 96,90%
Coeficiente de determinação ajustado (R^2 ajustado)	: 96,48%
Desvio padrão da estimativa	: 0,379278
Valor médio absoluto dos resíduos	: 0,290047
Teste estatístico Durbin-Watson..(p=0,1277)	: 2,34213

O valor significativo do coeficiente de determinação, mostrado acima, indica que o modelo ajustado explica mais de 96% da variabilidade do custo da rede coletora.

Entretanto, é preciso que se verifique a existência de deficiências que possam mascarar esse resultado. A análise de resíduos permite detectar essas deficiências de ajuste do modelo, tais como:

- A média dos erros não é nula.
- A variância dos resíduos não é constante.

- Os resíduos não são normalmente distribuídos.
- Existem variáveis importantes excluídas do modelo.
- A função de regressão não é linear.
- Os termos de resíduos não são independentes.
- Existem observações estranhas (outliers).

Essas condições, em geral, podem ser detectadas pela análise visual dos resíduos plotados. A análise é feita basicamente pela visualização da distribuição dos resíduos plotados em relação aos valores estimados das variáveis do modelo. Logo, se o modelo é bom, não há nenhuma relação de tendência entre os resíduos e as variáveis que definem o modelo.

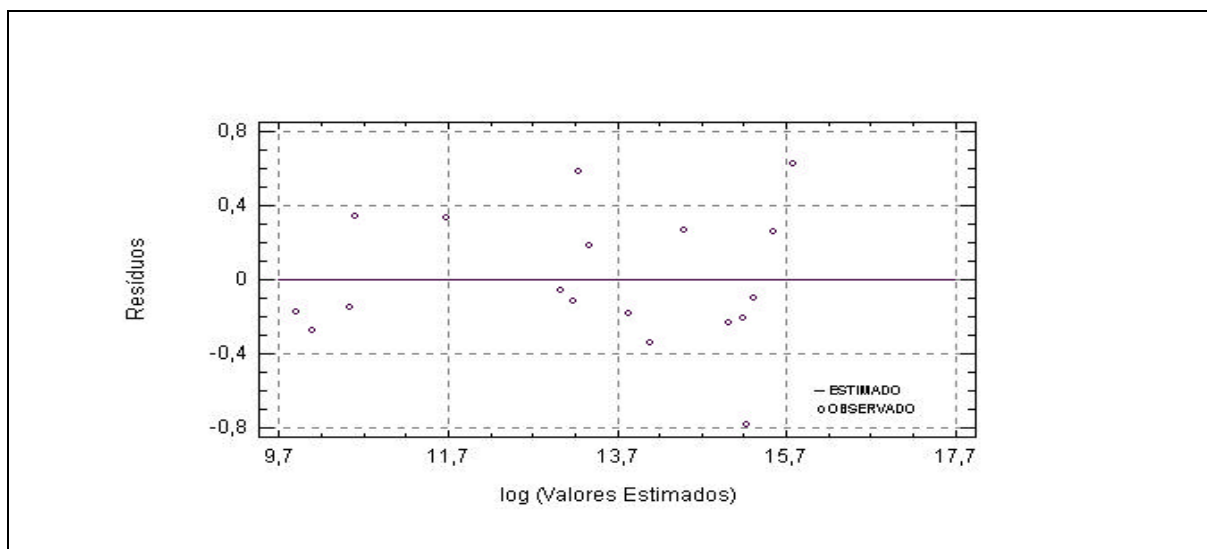


Figura 4.4. Resíduos dos custos estimados de redes coletoras.

No gráfico acima, onde os resíduos estão plotados em relação aos valores estimados, pode-se observar que os mesmos não apresentam uma distribuição que caracterize alguma tendenciosidade, que pode ser confirmada pelo teste de Durbin-Watson, cujo valor p é superior a 0,05, indicando que provavelmente não há autocorrelação nos resíduos.

Para confirmar a ausência de deficiências no modelo, são traçados também os gráficos de resíduos, em relação as variáveis independentes, área de abrangência e densidade populacional, os quais não apresentam nenhuma deficiência.

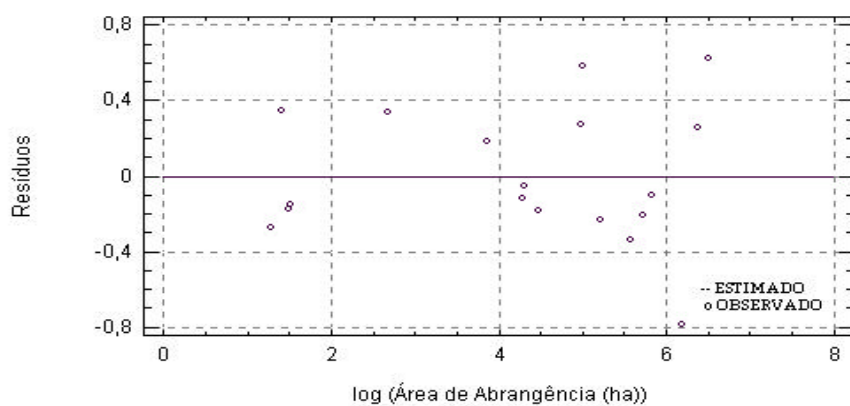


Figura 4.5. Resíduos dos custos estimados de rede coletora em relação à variável área de abrangência.

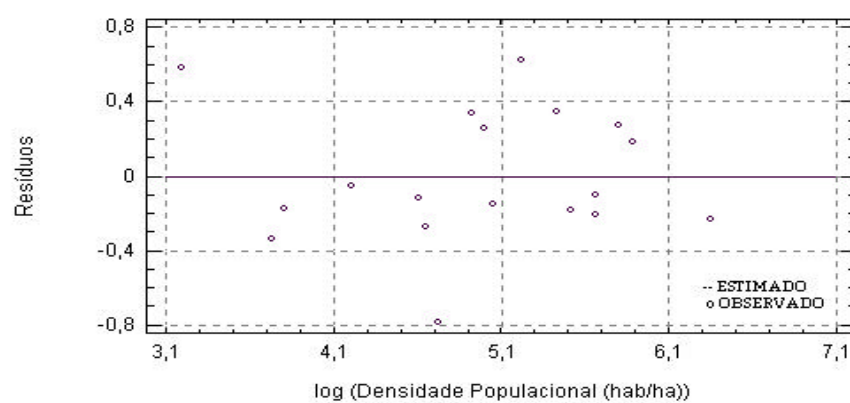


Figura 4.6. Resíduos dos custos estimados de rede coletora em relação à variável densidade populacional.

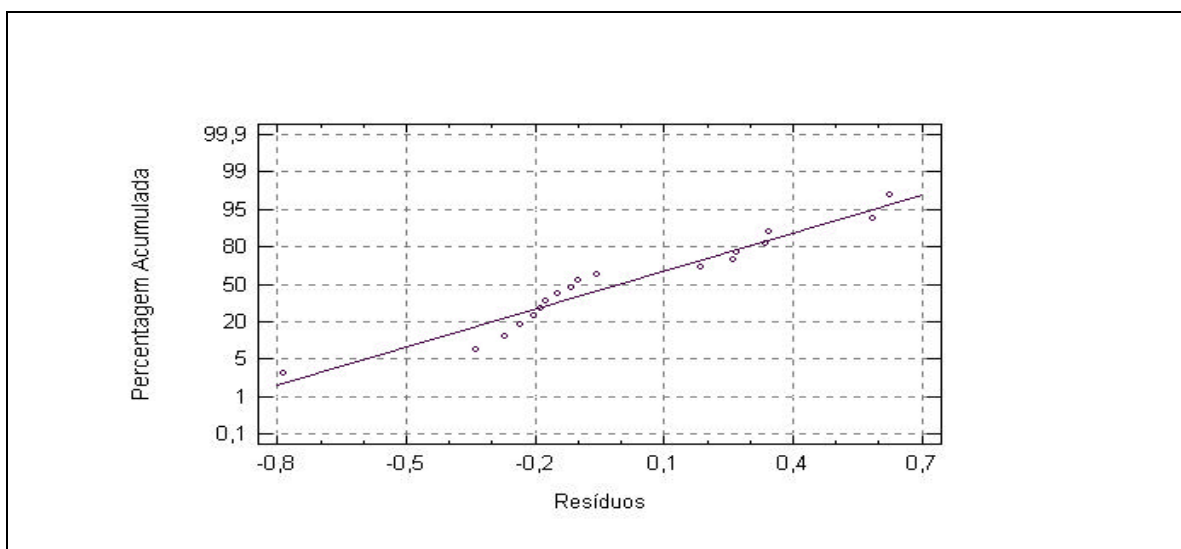


Figura 4.7. Gráfico da probabilidade normal dos resíduos dos custos estimados de redes coletoras.

Na figura 4.7. acima, onde foi traçada a probabilidade normal dos resíduos, verifica-se que o modelo analisado não apresenta indício de que haja deficiência quanto à normalidade da distribuição de probabilidade dos resíduos.

Confirma-se, assim, o que já foi observado na análise visual do gráfico de resíduos.

Para a maior precisão do modelo, pode-se retirar os dados do ponto, destacado com seta na figura 4.2, considerado espúrio (outlier) na análise dos resíduos e recalcular os valores estatísticos do modelo. O pequeno tamanho da amostra, apenas 18 pares de dados, não nos dá segurança que a retirada desse ponto traga significativa melhoria ao modelo. Basta observar a precisão do modelo no item 4.7. Validação dos Modelos.

Os gráficos a seguir mostram os valores estimados pelo modelo em relação aos apresentados nos dados da amostra.

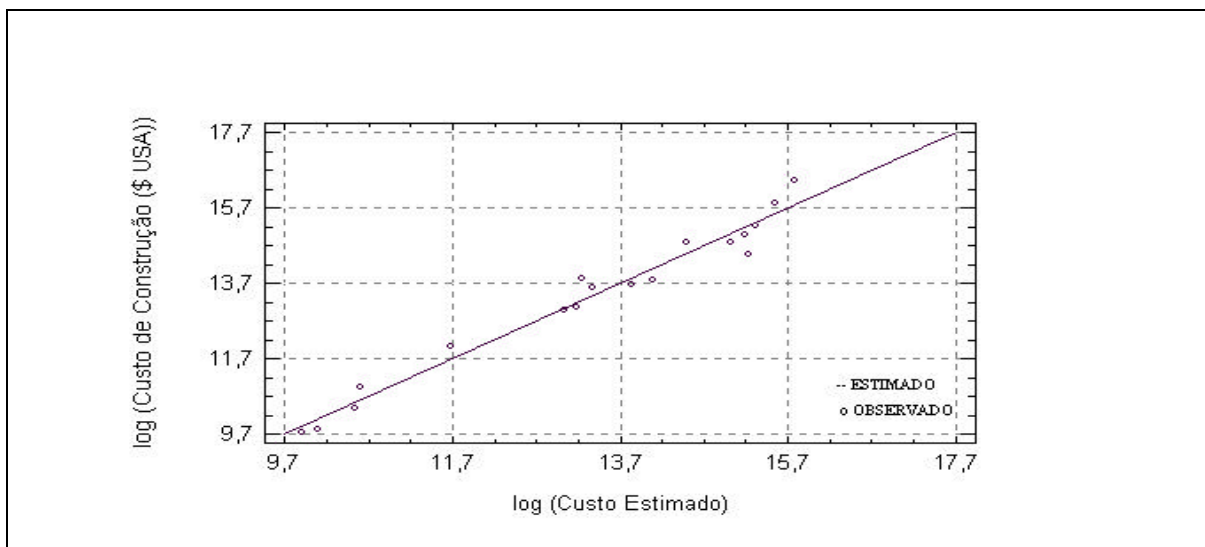


Figura 4.8. Custos estimados e observados de redes coletoras

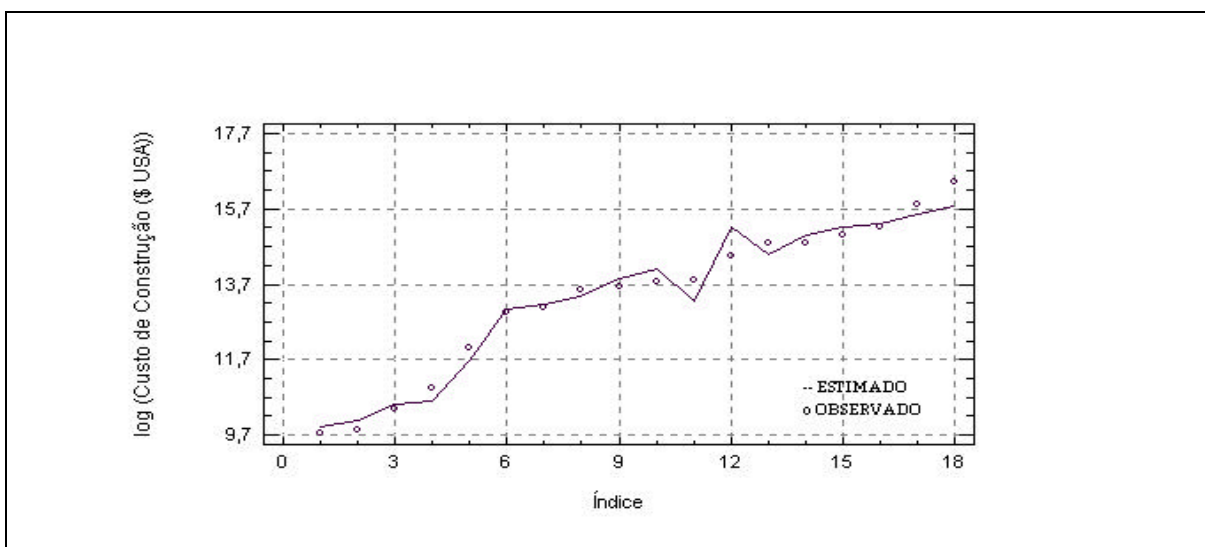


Figura 4.9. Custos estimados e observados de redes coletoras através de índice (número de ordem dos dados)

É possível perceber, nos gráficos apresentados anteriormente, que as variáveis área de abrangência e densidade populacional podem explicar a variação do custo das redes coletoras. Esse comportamento é o reflexo do alto valor obtido para a correlação das variáveis ao ajustar o modelo de custo.

Considerando que a regressão está concluída e testada estatisticamente, o modelo pode ser representado pela seguinte equação matemática:

$$Y_1 = 669.56 L_1^{0,4874} X_1^{1,0360} \quad (03)$$

4.2 Modelo de Custo de Ligação Predial de Esgoto Sanitário.

Implantada nos diâmetros de 100mm e 150mm, ela ocorre em grande quantidade no menor diâmetro(varia de 60 a 100%). Apenas nas áreas urbanas mais “verticalizadas” (onde há mais edifícios com mais de oito apartamentos) é que há um aumento significativo de ligações em 150mm.

O seu comprimento médio padronizado no projeto é de 8,0m. Cerca de 99% dos projetos de ligações usam material em PVC.

O custo da ligação compreende efetivamente: remoção e reposição de pavimentação, escavação e reaterro, aquisição e assentamento de tubos e conexões e caixa de inspeção com aquisição do tampão em ferro fundido.

Na tabela a seguir são apresentados os dados da amostra selecionada e normalizada para o ajuste do modelo de ligação predial. Estão apresentadas duas colunas de custos orçados, uma expressa em reais equivalentes ao mês de julho de 2000 [R\$(07/00)] e a outra em dólar americano (\$ USA). A função é ajustada pelo custo orçado em dólar.

O modelo de custo baseia-se em uma amostra de 23 dados históricos de projetos.

Tabela 4.4. Dados de projetos de ligações prediais.

Número de Ordem	Número de Ligações	Custo (R\$ 07/00)	Custo (\$ USA)	Custo (\$USA) por ligação
1	205	49.459,30	27.739,37	135,31
2	60	13.312,41	7.466,30	124,44
3	120	33.263,04	18.655,66	155,46
4	87	23.779,25	13.336,65	153,29
5	2.313	518.592,84	290.854,09	125,75
6	2.172	1.199.271,10	672.614,19	309,68
7	17.603	6.652.817,57	3.731.249,34	211,97
8	7.694	1.448.266,89	812.264,10	105,57
9	408	90.531,95	50.775,07	124,45
10	150	31.044,35	17.411,30	116,08
11	995	380.061,34	213.158,35	214,23
12	460	233.925,30	131.197,59	285,21
13	1.500	613.321,42	343.982,85	229,32
14	4.548	1.490.370,63	835.878,09	183,79
15	4.003	1.457.839,59	817.632,97	204,26
16	5.007	1.361.565,77	763.637,56	152,51
17	4.816	1.195.092,94	670.270,86	139,18
18	690	283.714,99	159.122,26	230,61
19	200	83.476,69	46.818,11	234,09
20	547	160.588,53	90.066,48	164,66
21	944	216.209,34	121.261,55	128,46
22	790	162.527,87	91.154,16	115,39
23	6.083	2.558.074,79	1.434.702,63	235,85

Fonte: Projetos arquivados na Companhia Catarinense de Águas e Saneamento (CASAN).

De posse dos dados da tabela 4.4., obtêm-se um gráfico que mostra o relacionamento entre os custos e o direcionador.

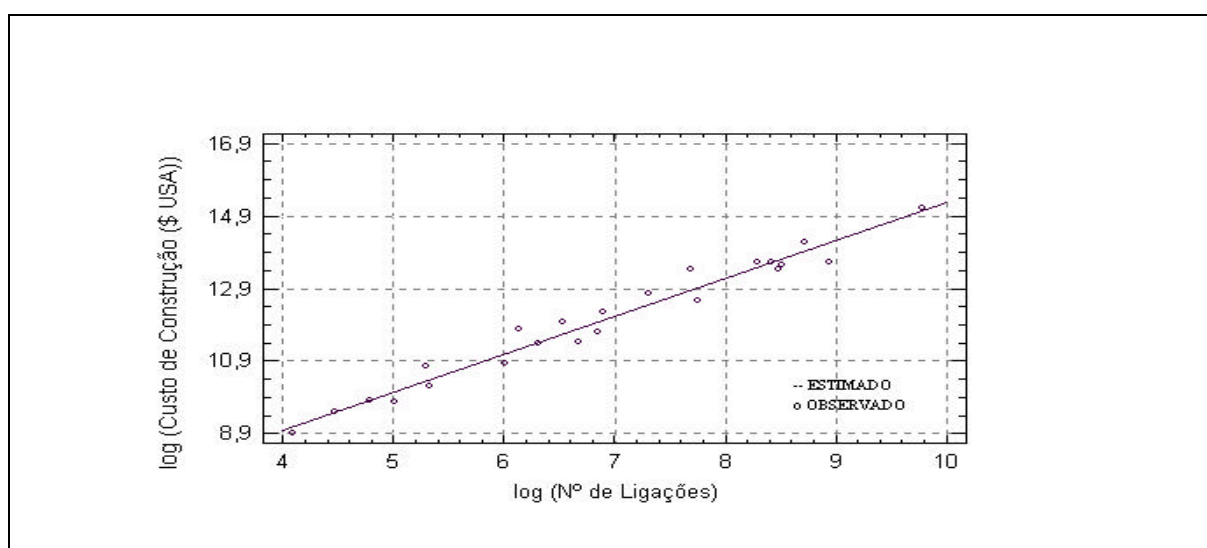


Figura 4.10. Custos de ligações prediais em função de número de ligações.

Depreende-se em relação aos pontos plotados no gráfico acima, que existe tendência de crescimento log-linear, podendo ter como resultante o traçado de uma reta. Porém, como se trata de uma relação logarítmica, o modelo de custo em função da variável número de ligações é não linear.

A equação matemática, utilizada para descrever a relação paramétrica do modelo com a variável explicativa número de ligações, é do tipo não linear, conforme mostrada abaixo:

$$Y_2 = K_2 L_2^a \quad (4)$$

Onde:

Y_2 = Custo de Ligação Predial de Esgoto Sanitário (\$ USA);

K_2 = Constante de Custo para Ligação Predial de Esgoto Sanitário;

L_2 = N.º de Ligações Prediais;

a = Expoente de L_2 que representa o Fator de Economia de Escala da Ligação Predial.

Para definição do modelo, é utilizada a regressão linear múltipla da transformada da equação (4):

$$\log(Y_2) = \log(K_2) + a \cdot \log(L_2) \quad (5)$$

As tabelas a seguir mostram os resultados finais da regressão para o ajustamento do modelo

Tabela 4.5. Ajuste do modelo de custo de ligações prediais.

Parâmetros	Estimativa	Erro Padrão	t Estatístico	Nível de Significância
$\log(K_4)$	4,85122	0,29823	16,2667	0,0000
a	1.04053	0,04227	24,6144	0,0000

O teste t de Student, acima, demonstra que o fator de economia de escala (expoente) é significativamente diferente de zero e que a variável número de ligações prediais é relevante na avaliação de custo e que se relacionam positivamente.

O erro padrão, por seu lado, medindo a precisão do fator de escala, mostra uma variação de apenas 4,2% em torno da média.

Tabela 4.6. Análise de Variância de Ajuste do Modelo.

Fonte	Somatório dos Quadrados	Graus de Liberdade	Quadrado Médio	Razão "F"	Valor "p"
Modelo	60,012	1	60,012	605,87	0,0000
Resíduo	2,080	21	0,0995		
Total	62,092	22			

A hipótese testada mostra que a Razão F é significativamente maior que um num nível de significância de 5%, indicando que o modelo tem boa aderência a tendência dos dados observados.

Outros dados estatísticos para o modelo:

Tamanho da amostra	23
Dados desconsiderados	0
Coeficiente de determinação (R^2)	96,65%
Coeficiente de determinação ajustado (R^2 ajustado)	96,49%
Desvio padrão da estimativa	0,3147
Valor médio absoluto dos resíduos	0,251
Teste estatístico Durbin-Watson ($p= 0,1303$)	1,598

Na tabela 4.7., a seguir, é mostrada a significância da variável independente no modelo. Ela tem participação relevante, se considerado que o Fator F é significativamente maior que um.

Tabela 4.7. Análise da variância das variáveis independentes.

Fonte	Somatório dos Quadrados	Graus de Liberdade	Quadrado Médio	Razão "F"	Valor "p"
log(L ₄)	60,012	1	60,012	605,87	0,0000
Modelo	60,012	1			
		2			

O valor significativo do coeficiente de determinação, mostrado anterior, indica que o modelo ajustado explica mais de 96% da variabilidade do custo da ligação predial.

A análise de resíduos, através da observação dos gráficos de resíduos adiante, permitirá detectar se há deficiências de ajuste do modelo.

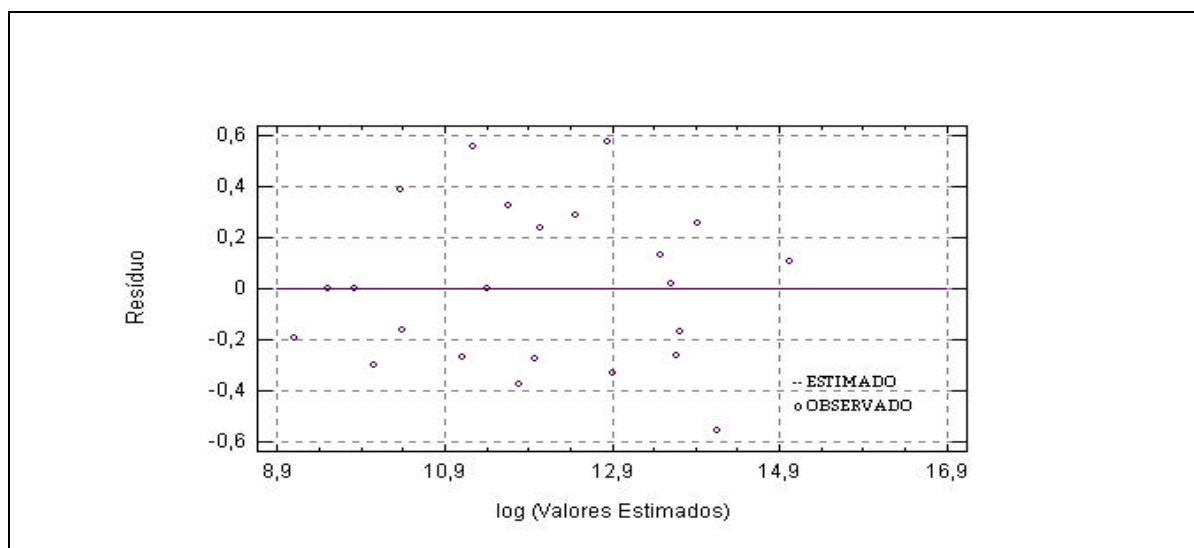


Figura 4.11. Resíduos dos custos estimados de ligações prediais.

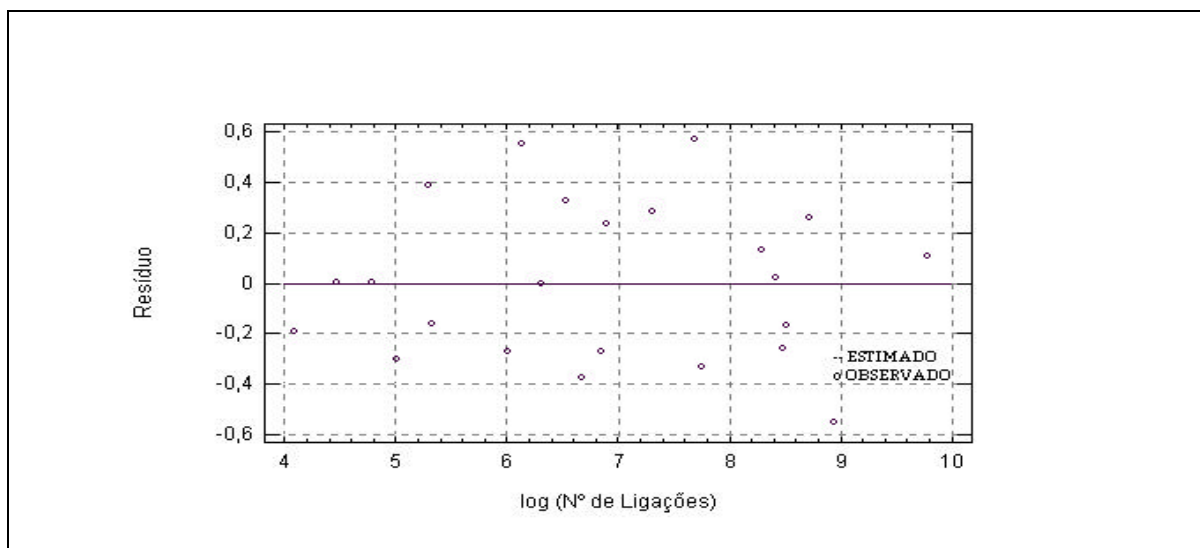


Figura 4.12. Resíduos dos custos estimados relacionados à variável n.º de ligações.

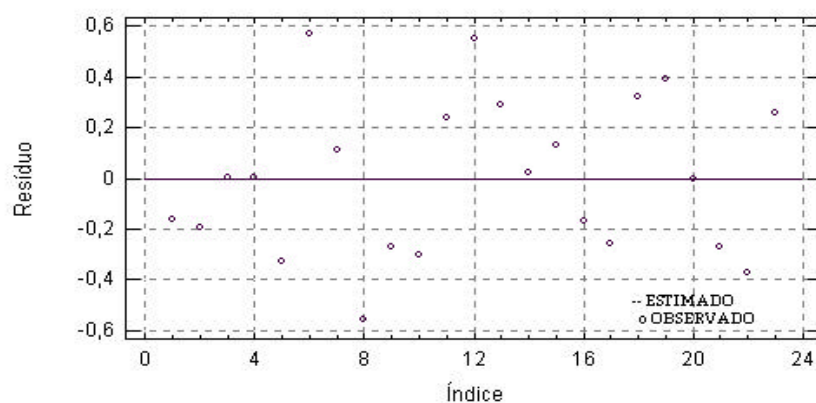


Figura 4.13. Resíduos em função de índice (n.º de ordem dos dados)

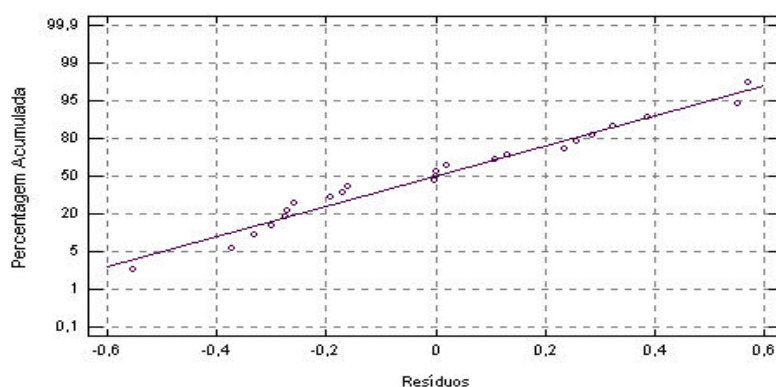


Figura 4.14. Gráfico de probabilidade normal dos resíduos dos custos estimados de ligações prediais.

Ao se observar os gráficos de resíduos, nas quatro figuras anteriores, é possível perceber que os mesmos não apresentam uma distribuição que caracterize alguma tendenciosidade, tampouco há autocorrelação bem como confirmam que a variância é constante.

Na figura 4.14., o gráfico da probabilidade normal dos resíduos mostra que não há deficiência quanto a normalidade de sua distribuição.

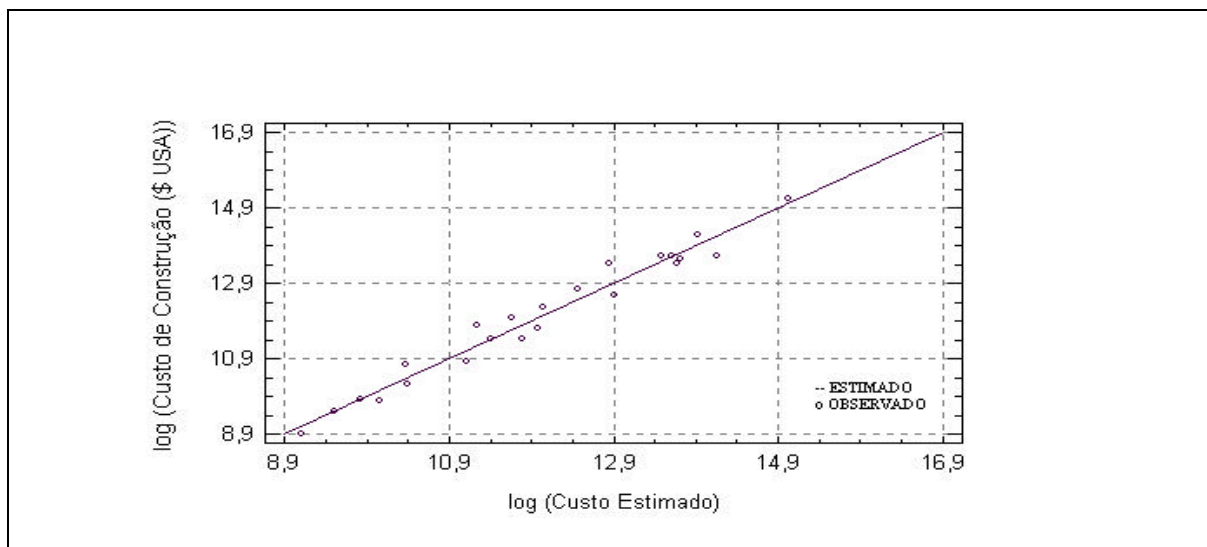


Figura 4.15. Custos estimados e observados de ligações prediais.

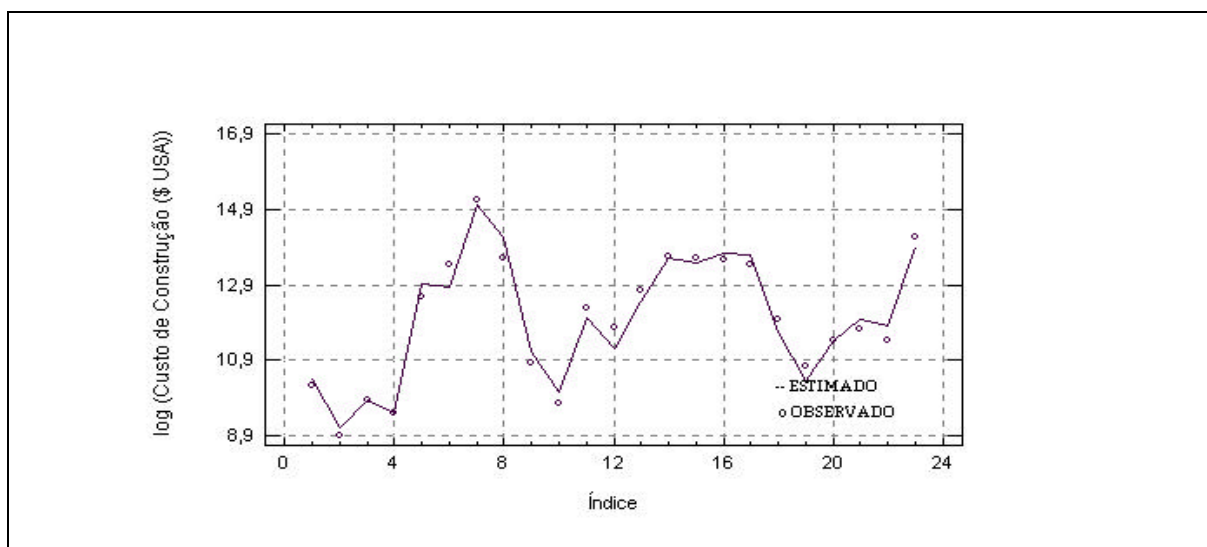


Figura 4.16. Custos estimados e observados de ligações prediais através de índice (n.º de ordem dos dados).

Os dois gráficos acima confirmam que a variável, número de ligações, explica a variação dos custos de ligações prediais de esgoto sanitário. Isso reflete o alto coeficiente de determinação do modelo.

Considerando que o modelo está concluído e testado estatisticamente, ele está representado abaixo pela transformada e pela forma direta da equação matemática.

$$\log(Y_2) = 4,85122 + 1,04053.\log(L_2) \quad (6)$$

Para facilitar o uso da equação, ela é revertida a forma direta:

$$Y_2 = 127,89L_2^{1,040} \quad (7)$$

4.3 Modelo de Custo de Estação Elevatória de Esgoto Sanitário

Seus custos expressam os gastos com os insumos, mão-de-obra, materiais e equipamentos, necessários à implantação de uma unidade de estação elevatória, cujos os conjuntos motobombas são em sua maioria submersíveis.

Salvo pequenas alterações, essas estações são padronizadas. Elas compreendem um poço circular em concreto com profundidade de 4,00 a 8,00m onde ficam alojados os conjuntos motobombas; um poço de gradeamento; as tubulações de sucção e do barrilete; a estrutura de içamento das motobombas e a estrutura do quadro de comando.

Os referidos custos incluem, escavação, escoramento, esgotamento e rebaixamento do lençol freático, reaterro, concreto em geral, revestimentos, impermeabilização, instalações hidráulicas, aquisição e montagem de equipamentos elétricos e eletromecânicos, tubulações e testes de partida.

É possível observar adiante, que a melhor relação com o custo envolve a combinação da capacidade de recalque (Q) com a altura manométrica (H).

O modelo da equação de custo estudada para estação elevatória de esgoto sanitário está fundamentado nas duas variáveis independentes citadas acima. A primeira, identificada pela capacidade de recalque de esgoto medida em litros por segundo (l/s), que é dada pela vazão de projeto, e a segunda, a altura manométrica (H.m.) medida em metros de coluna d'água (m.c.a.).

Na tabela a seguir são apresentados os dados da amostra obtida para o ajustamento do modelo.

Tabela 4.8. Dados de projetos de estações elevatórias de esgoto sanitário

Número De Ordem	Capacidade (l/s)	Altura Manométrica. (m.c.a)	Custo (R\$ 07/00)	Custo (\$ USA)
1	0,31	12,36	5.284,48	2.963,81
2	0,56	7,98	5.284,48	2.963,81
3	0,79	6,45	5.284,48	2.963,81
4	1,65	4,46	23.831,3	13.365,84
5	3,33	31,00	40.131,74	22.507,99
6	8,33	11,05	42.589,92	23.886,66
7	1,74	35,00	47.071,90	26.400,39
8	7,76	5,00	52.005,26	29.167,28
9	19,17	6,00	57.779,15	32.405,58
10	8,69	10,00	72.847,50	40.856,70
11	5,92	24,70	77.269,31	43.336,69
12	6,00	26,35	82.150,14	46.074,11
13	24,48	8,00	91.382,13	51.251,90
14	36,61	23,00	103.738,17	58.181,81
15	27,66	12,00	131.462,53	73.731,09
16	13,10	26,63	138.420,86	77.633,68
17	44,00	18,50	178.381,13	100.045,50
18	39,56	27,37	181.814,73	100.864,88
19	64,00	17,48	181.814,73	101.971,25
20	26,46	13,22	192.567,06	108.001,72
21	96,00	9,66	214.458,27	120.279,46
22	65,77	8,00	215.380,39	120.796,63
23	78,00	10,00	235.251,37	131.941,32
24	46,00	10,00	249.818,49	140.111,32
25	98,00	6,07	262.102,30	147.000,73
26	75,00	8,00	296.367,12	166.218,24
27	85,00	12,00	308.348,58	172.938,07
28	73,00	10,41	357.783,11	200.663,55
29	110,00	10,93	409.818,28	229.847,61
30	176,72	17,68	421.707,13	236.515,50
31	236,00	12,00	515.594,57	289.172,50
32	204,00	39,00	653.189,46	366.342,94
33	263,00	25,26	763.531,65	428.228,63

Fonte: Projetos Arquivados na Companhia Catarinense de Águas e Saneamento(CASAN).

Com o objetivo de definir o modelo de custo, foi verificada a tendência inicial do comportamento da amostra. Uma variável de configuração (altura manométrica) e outra de capacidade (vazão) são as selecionadas para direcionar a variável de custo.

Os dados estão relacionados nos gráficos plotados nas figuras 4.17, 4.18. e 4.19., a seguir, onde pode-se fazer uma análise inicial do comportamento dessas variáveis.

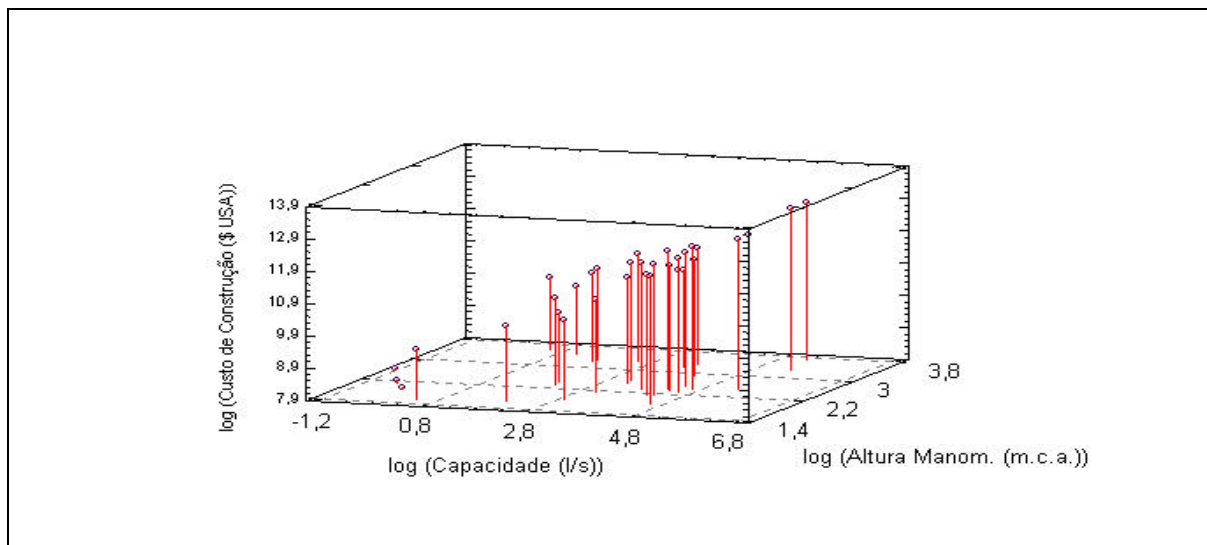


Figura 4.17. Custos de estações elevatórias em função de capacidade e altura manométrica.

Observando o gráfico acima, depreende-se que a medida que aumenta a capacidade o custo também aumenta. De forma idêntica, embora não tão expressiva, ocorre também em relação a variável de configuração.

Essas tendências flagrantes mostram que as variáveis escolhidas para explicar os custos de construção de estações elevatórias apresentam evidências de correlação de forma singular ou conjunta.

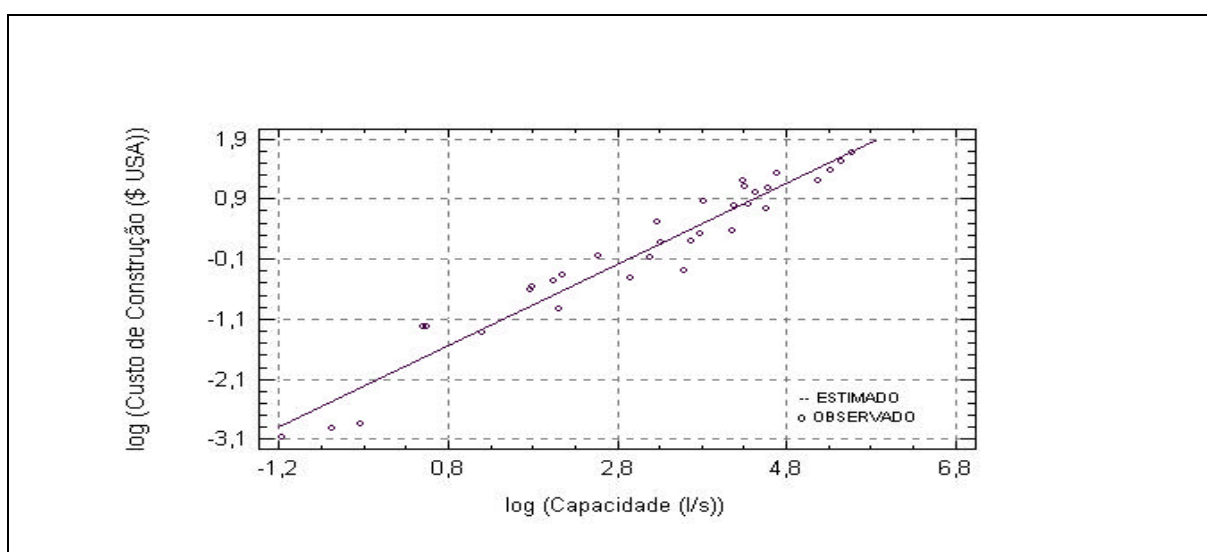


Figura 4.18. Custos de estações elevatórias em função de capacidade.

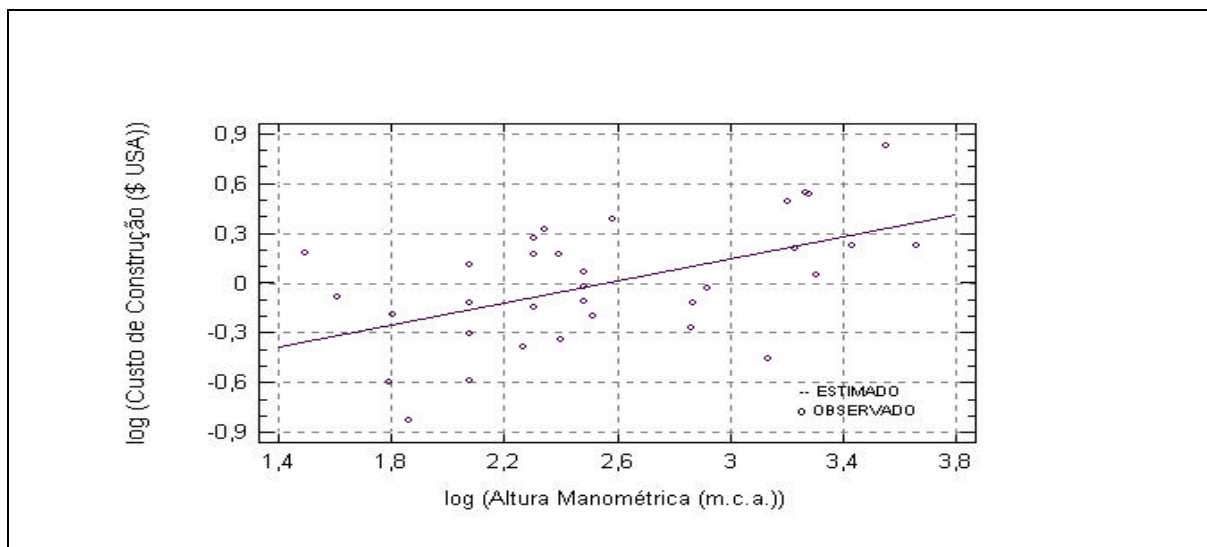


Figura 4.19. Custos de estações elevatórias em função de altura manométrica.

Analisando os gráficos das figuras 4.18. e 4.19., constata-se igualmente as tendências já observadas na figura 4.17. A variável capacidade mostra uma melhor relação com os custos, se comparada com a variável altura manométrica, pois os dados observados pouco se dispersam em relação a reta log-linear.

Apresenta-se a seguir a equação matemática que melhor expressa a relação e seu consequente modelo paramétrico:

$$Y_3 = K_3 L_3^b X_3^a \quad (8)$$

Onde:

Y_3 = custo de estação elevatória de esgoto sanitário (\$ USA);

K_3 = constante de custo para estação elevatória;

L_3 = altura manométrica de recalque da estação em metros de coluna de água (m.c.a.);

b = expoente de L_3 que representa o fator de economia de escala da altura manométrica;

X_3 = capacidade da estação elevatória (vazão de projeto) (l/s);

a = expoente de X_3 que representa o fator de economia de escala da capacidade da estação elevatória.

A regressão linear múltipla utiliza a transformada da equação, mostrada abaixo, para desenvolver o ajuste do modelo da relação paramétrica de custo.

$$\log(Y_3) = \log(K_3) + b.\log(L_3) + a.\log(X_3) \quad (9)$$

A tabela a seguir mostra os resultados estatísticos da regressão múltipla de ajustamento do modelo.

Tabela 4.9. Ajuste do modelo de custo de estações elevatórias.

Parâmetros	Estimativa	Erro Padrão	t estatístico	Nível de Significância
K_3	8,1329	0,25983	31,3006	0,0000
a	0,6793	0,03144	21,6016	0,0000
b	0,3297	0,09645	3,4183	0,0018

Os valores dos testes, na tabela acima, mostram que os fatores de economia de escala (expoentes) são significativamente diferentes de zero. Isso é, ambas as variáveis independentes, capacidade e altura manométrica, são relevantes na avaliação de custo de estação elevatória de esgoto sanitário e se correlacionam positivamente com a variável dependente.

Os testes de significância da hipótese nula das variáveis direcionadoras, analisados pela distribuição de Student, estão representados nas duas últimas colunas da tabela 4.9 com nível de significância nulo para o expoente a e próximo da nulidade para o expoente b . Isto indica que a probabilidade dos fatores de escala dessas variáveis, definidos pela correlação, serem zero é igualmente nula. Se o expoente de uma variável é zero, qualquer valor que ela assuma terá como resultado sempre o número um. Significando que a variável não influenciará os valores de custo resultantes do modelo.

Os resultados do erro padrão, cujos valores medem a precisão dos fatores de escala estimados em termos absolutos, mostram uma variação maior para o fator de escala da variável altura manométrica. Esse resultado demonstra menor precisão na determinação desse fator. Contudo, em relação ao fator de escala da variável capacidade, o erro padrão é menor e representa uma variação de apenas 4,6% em torno da média.

Para a variância, estão representados os resultados da análise completa na tabela a seguir.

Tabela 4.10. Análise da Variância do Ajuste do Modelo.

Fonte	Somatório dos quadrados	Graus de liberdade	Quadrado médio	Razão F	Valor p
Modelo	51,2364	2	25,6182	252,13	0,0000
Resíduos	3,0482	30	0,1016		
Total	54,2846	32			

O teste de hipótese pela distribuição F de Fisher realizado para o modelo de custo, onde a Razão F é significativamente maior do que um, levou a rejeição da hipótese nula para um nível de significância de 5%. Isso significa que o modelo tem boa aderência à tendência dos dados observados.

Esse teste baseia-se na relação entre o quadrado médio devido à regressão e o quadrado médio residual para explicar a aceitação ou rejeição da hipótese nula. Quanto maior o F observado maior é a probabilidade de rejeição da hipótese nula.

Outros dados estatísticos para o modelo são apresentados abaixo:

Número de dados analisados	: 33
Número de dados desconsiderados	: 0
Coeficiente de determinação (R^2)	: 94,38%
Coeficiente de determinação ajustado (R^2 ajustado)	: 94,01%
Desvio padrão da estimativa	: 0,318758
Valor médio absoluto dos resíduos	: 0,255649
Teste estatístico Durbin-Watson..(p=0,3968)	: 2,01762

Elaboradas as análises de resíduos para confirmar a qualidade da relação paramétrica (modelo) de custo de estação elevatória, os resultados mostram que a equação matemática é satisfatória e que a capacidade e a altura manométrica explicam 94% das variações de custo.

Nas seis figuras a seguir estão plotados os custos estimados pelo modelo em relação aos custos observados e as demais relações dos resíduos que explicam o bom ajuste do modelo.

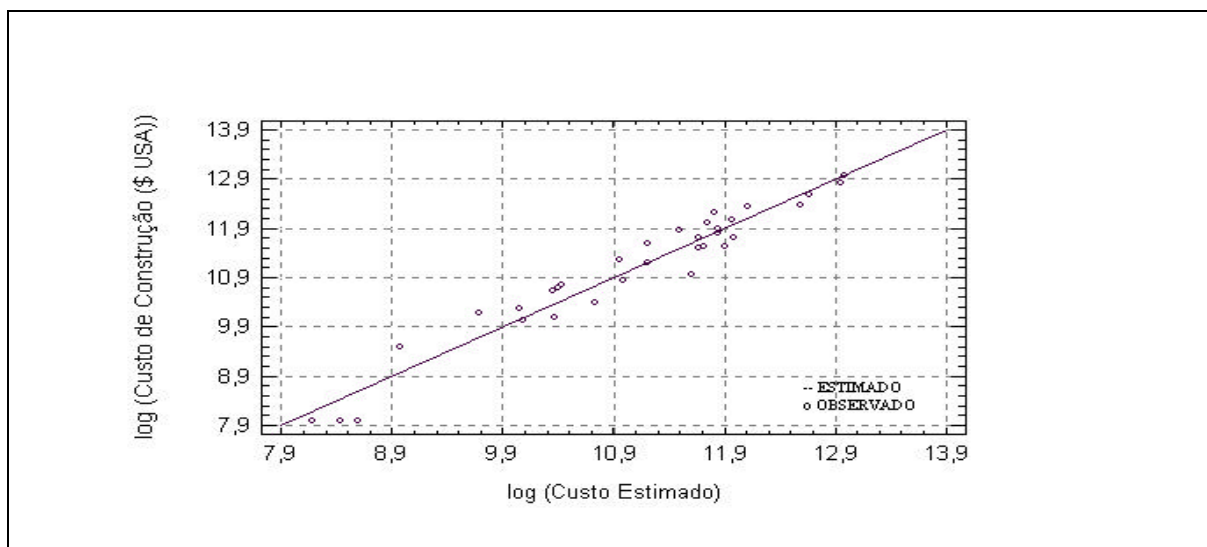


Figura 4.20. Custos estimados e respectivos custos observados para estações elevatórias.

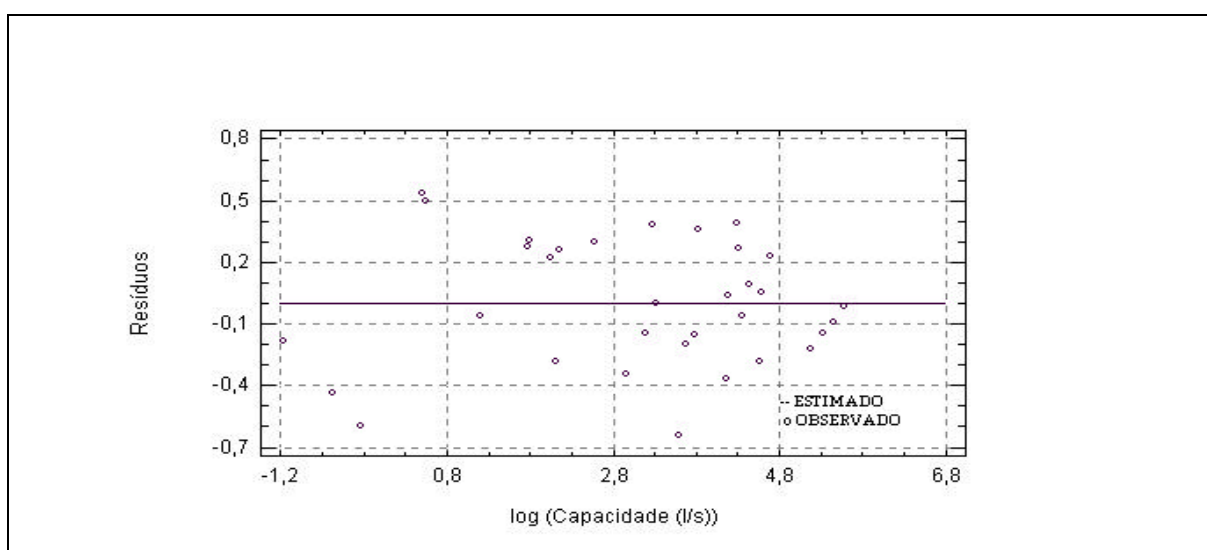


Figura 4.21. Resíduos dos custos estimados de estações elevatórias relacionados com a variável capacidade.

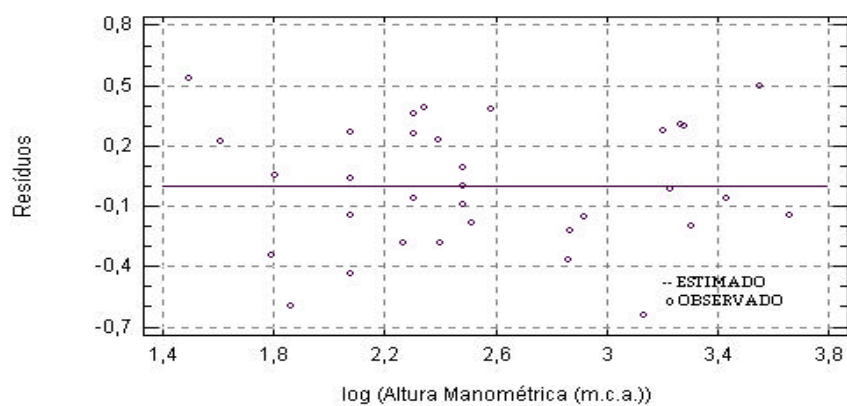


Figura 4.22. Resíduos dos custos estimados de estações elevatórias em relação à variável altura manométrica.

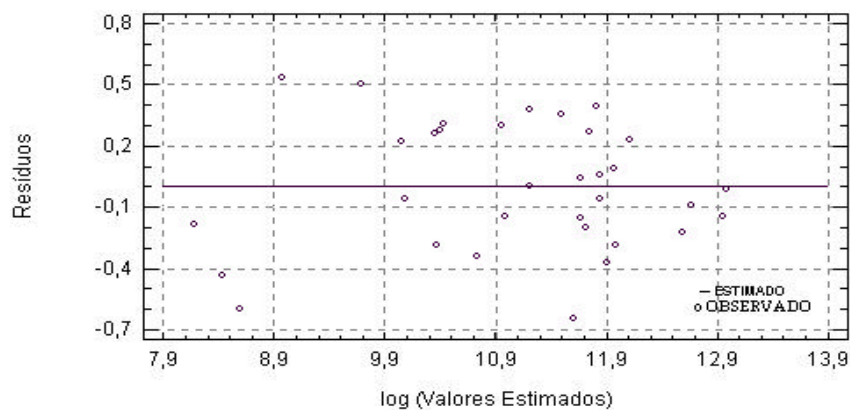


Figura 4.23. Resíduos dos custos estimados de estações elevatórias.

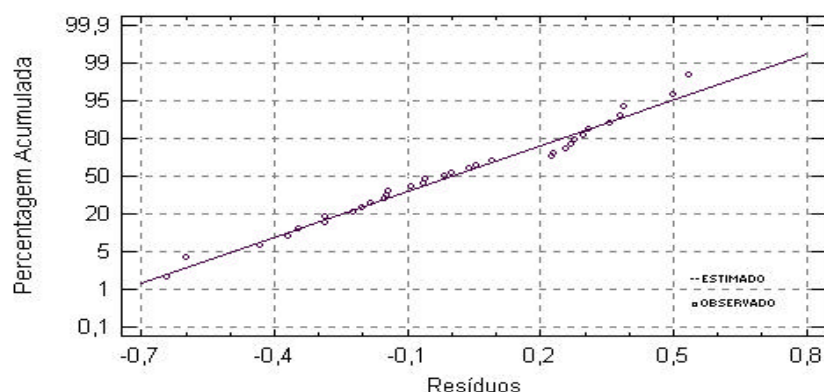


Figura 4.24. Gráfico da probabilidade normal dos resíduos dos custos estimados das estações elevatórias.

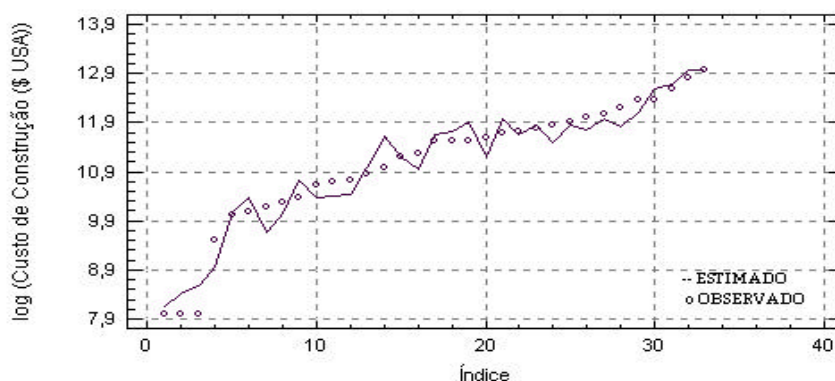


Figura 4.25. Custos estimados e observados de estações elevatórias através de índice (número de ordem dos dados)

Verificadas as deficiências no modelo pela análise dos resíduos, como não há evidências delas, ele está definitivamente concluído e resulta na equação abaixo:

$$Y_3 = 3,29 \times 10^3 L_3^{0,330} X_3^{0,679} \quad (10)$$

4.4 Modelo de Custo de Emissário de Esgoto Sanitário.

Seu custo é caracterizado pelo seguintes serviços: locação e cadastro, limpeza do terreno, remoção e reposição de pavimentação, escavação, reaterro, substituição

de solos, escoramentos, esgotamento e rebaixamento do lençol freático, embasamentos, aquisição e assentamento de tubos, peças e conexões e testes de estanqueidade.

Vários dados de projeto de emissário, relacionáveis ao seu custo, foram selecionados. Todos são considerados direcionadores de custo, pois é fácil entender que cada um deles ou em conjunto explicam a variação do custo do emissário.. Alguns desses parâmetros são mais simples e de mais baixo custo para obtê-los.

A vazão de projeto e a extensão são dados básicos e obtidos na fase de estudo do emissário.

Desenvolvendo os testes de significância, é possível observar que a melhor relação com o custo é a combinação que envolve a vazão de projeto e a extensão de emissário.

O modelo de custo, definido para essa parte de sistema é, portanto, explicado pelas variáveis independentes: vazão de projeto medida em litros por segundo (l/s) e extensão de emissário medida em metros (m).

O banco de dados históricos normalizado e consistente para formulação da relação paramétrica de custo do emissário, é mostrado na tabela a seguir.

Tabela 4.11. Dados de projetos de emissários.

Número de Ordem	Vazão (l/s)	Extensão (m)	Custo (R\$ 07/00)	Custo (\$ USA)
1	0,11	75,00	631,91	354,41
2	0,31	95,00	875,85	491,22
3	0,56	107,00	1.025,10	574,93
4	0,79	58,00	520,15	291,73
5	3,33	725,00	47.418,14	26.594,58
6	6,00	352,00	25.767,79	14.451,93
7	9,66	280,00	56.738,38	31.821,86
8	8,69	760,00	86.167,04	48.327,00
9	6,24	568,00	40.549,18	22.742,11
10	13,10	1.100,00	166.888,98	93.600,10
11	18,00	438,00	37.574,03	21.073,49
12	31,00	526,00	64.616,76	36.240,47
13	26,46	1.190,00	216.429,46	121.385,00
14	64,00	640,00	123.455,63	69.240,40
15	18,00	1.003,00	113.054,75	63.407,04
16	73,00	292,00	54.315,42	30.462,94
17	65,77	2.534,00	560.893,84	314.578,71
18	75,00	108,00	23.977,45	13.447,81
19	46,00	268,40	83.675,99	46.929,89
20	39,56	2.380,00	322.162,53	180.685,66
21	97,50	294,00	46.904,01	26.306,23
22	110,00	928,00	182.666,91	102.449,19
23	78,00	1.717,00	383.508,61	215.091,76
24	27,66	979,00	275.842,92	154.707,19
25	85,00	240,00	101.572,54	56.967,21
26	236,00	817,90	371.112,39	208.139,31
27	263,00	3.515,00	893.418,28	501.075,87
28	204,00	1.352,00	651.267,99	365.265,28
29	236,00	1.348,40	913.499,97	512.338,74
30	176,72	1.684,00	746.928,55	418.916,74
31	236,00	239,10	172.813,44	96.922,85

Fonte: Projetos da Companhia Catarinense de Águas e Saneamento (CASAN).

Com o objetivo de definir o modelo de custo, foi verificada a tendência inicial do comportamento da amostra. Uma variável de configuração (extensão) e outra de capacidade (vazão) são as selecionadas para direcionar a variável de custo.

Os dados estão relacionados nos gráficos plotados nas figuras 4.26, 4.27. e 4.28., a seguir, onde pode-se fazer uma análise inicial do comportamento dessas variáveis.

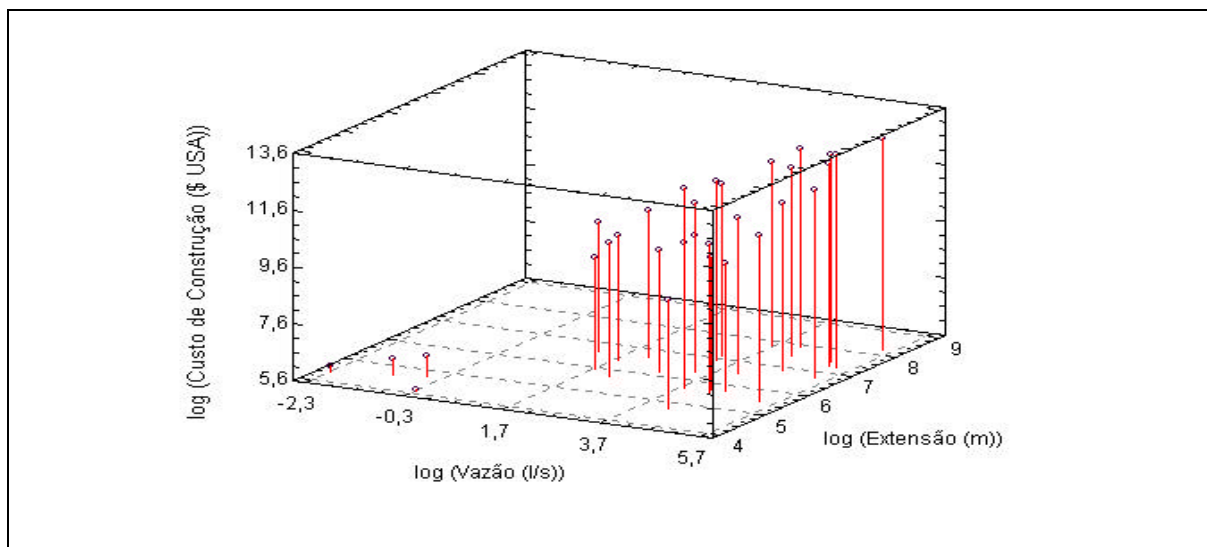


Figura 4.26. Custos de emissários em função de vazão e extensão.

Observando o gráfico acima, depreende-se que a medida que aumenta a vazão o custo também aumenta. De forma idêntica ocorre também em relação a variável de configuração.

Essas tendências flagrantes mostram que as variáveis escolhidas para explicar os custos de construção de emissários apresentam evidências de correlação de forma singular ou conjunta.

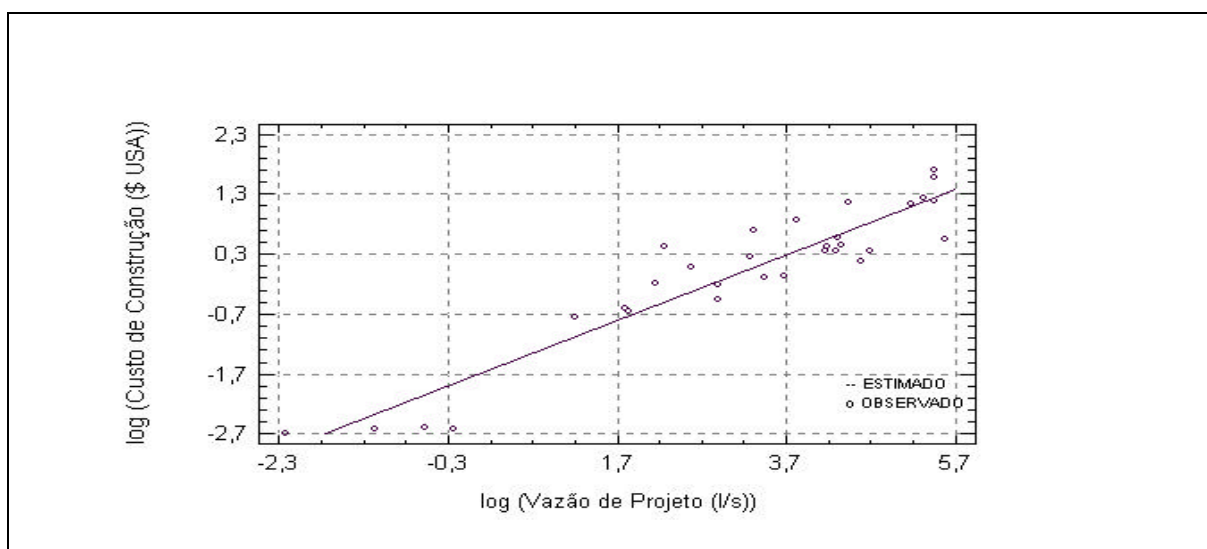


Figura 4.27. Custos de emissários em função de vazão de projeto.

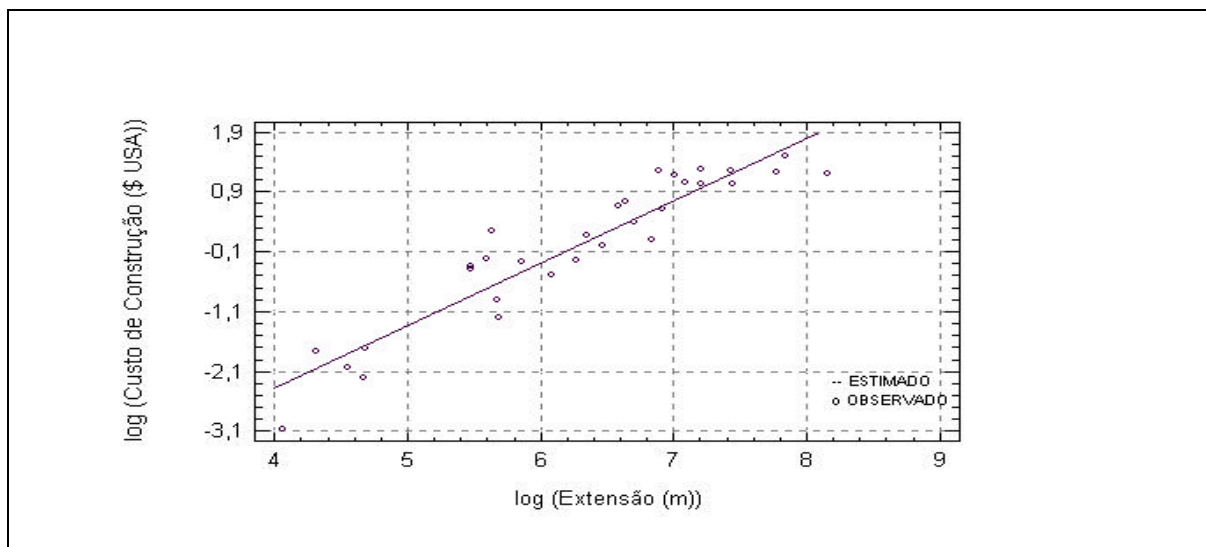


Figura 4.28. Custos de emissários em função de sua extensão.

Analisando os gráficos das figuras 4.27. e 4.28. constata-se igualmente as tendências já observadas na figura 4.26. As variáveis explanadoras mostram uma boa relação com os custos pois os dados observados pouco se dispersam em relação a reta log-linear.

A expressão matemática definida para representar o modelo é a seguinte:

$$Y_4 = K_4 L_4^b X_4^a \quad (11)$$

Onde:

Y_4 = Custo de Emissário de Esgoto Sanitário (\$ USA);

K_4 = Constante de Custo para Emissário de Esgoto Sanitário;

L_4 = Extensão do Emissário (m);

b = Expoente de L_4 que representa o Fator de Economia de Escala da Extensão;

X_4 = Vazão de Projeto do Emissário (l/s);

a = Expoente de X_4 que representa o Fator de Economia de Escala da Vazão de Projeto.

Usando regressão linear múltipla, na transformada da equação, chegou-se aos resultados finais de ajustamento da relação paramétrica de custo, conforme mostrado na tabela 4.12.

Equação transformada:

$$\log(Y_4) = \log(K_4) + b \cdot \log(L_4) + a \cdot \log(X_4) \quad (12)$$

Tabela 4.12. Ajuste do modelo de custo de emissários.

Parâmetros	Estimativa	Erro Padrão	t Estatístico	Nível de Significância
log(K ₄)	2,3177	0,50042	4,63141	0,0001
a	0,5483	0,04843	11,3225	0,0000
b	1,0435	0,09160	11,3922	0,0000

O teste t estatístico, na tabela acima, mostra que os fatores de economia de escala (expoentes) e a constante são significativamente diferentes de zero. Ambas as variáveis direcionadoras: vazão de projeto e extensão do emissário são relevantes na avaliação de custo de emissário e se correlacionam positivamente com a variável dependente.

Considerando que os níveis de significância são praticamente nulos, indicam que a probabilidade dos parâmetros da expressão do modelo serem zero é igualmente nula.

Os resultados do erro padrão que indicam a precisão dos fatores de escala, mostram uma variação praticamente igual, de apenas 8,7% em torno da média.

A análise de variância do modelo é mostrada adiante, na tabela 4.13.

Tabela 4.13. Análise da variância do ajuste do modelo.

Fonte	Somatório dos Quadrados	Graus de Liberdade	Quadrado Médio	Razão F	Valor p
Modelo	123,761	2	61,8805	339,39	0,0000
Resíduo	5,105	28	0,1823		
Total	128,866	30			

A hipótese testada mostra que a Razão F é significativamente maior que um dentro do nível de significância de 5%. Isso significa que o modelo tem boa aderência a tendência dos dados observados.

Outros dados estatísticos para o modelo:

Tamanho da amostra	31
Dados desconsiderados	0
Coeficiente de determinação (R^2)	96,04%
Coeficiente de determinação ajustado (R^2 ajustado)	95,76%
Desvio padrão da estimativa	0,427
Valor médio absoluto dos resíduos	0,3367
Teste estatístico Durbin-Watson ($p= 0,037$)	1,4308

Na tabela 4.14., a seguir, é mostrada a significância de cada variável independente no modelo. Elas têm participação relevante, se considerado que o Fator F é significativamente maior que um.

Tabela 4.14. Análise da variância das variáveis independentes.

Fonte	Somatório dos Quadrados	Graus de Liberdade	Quadrado Médio	Razão F	Valor p
$\log(X_4)$	100,098	1	100,098	549,00	0,0000
$\log(L_4)$	23,663	1	23,663	129,78	0,0000
Modelo	123,761	2			

O coeficiente de determinação, citado acima, indica que o modelo explana pelo menos 95% da variabilidade dos custos (variável dependente).

Para verificar deficiências, que possam mascarar o resultado desse coeficiente serão analisados visualmente os resíduos nos gráficos a seguir apresentados.

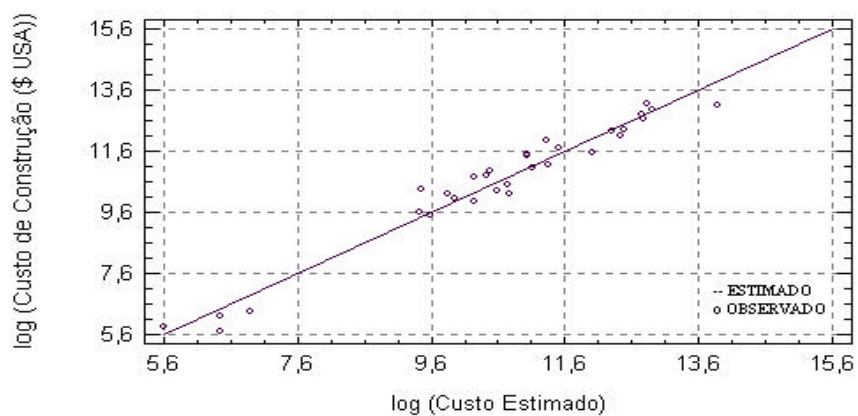


Figura 4.29. Resíduos dos custos estimados de emissários.

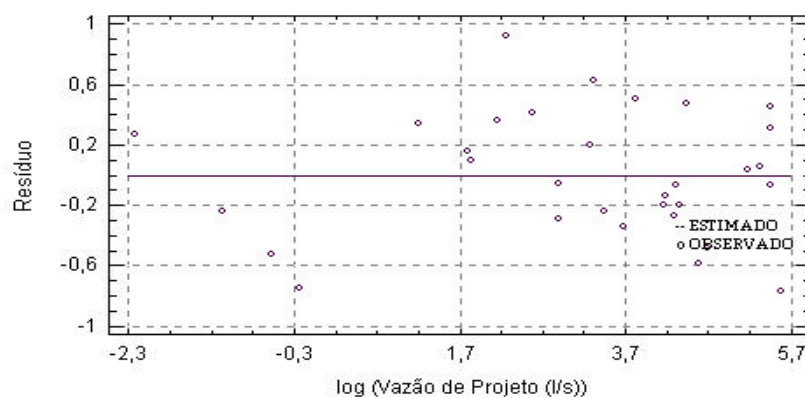


Figura 4.30. Resíduos dos custos estimados de emissários em relação à vazão de projeto.

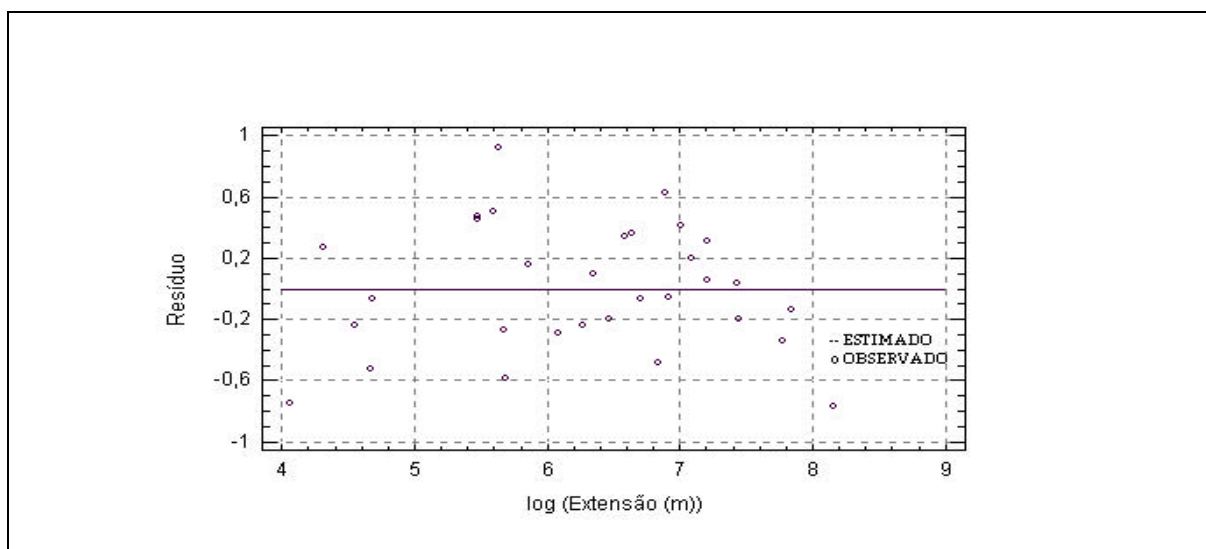


Figura 4.31. Resíduos dos custos estimados de emissários em relação à extensão

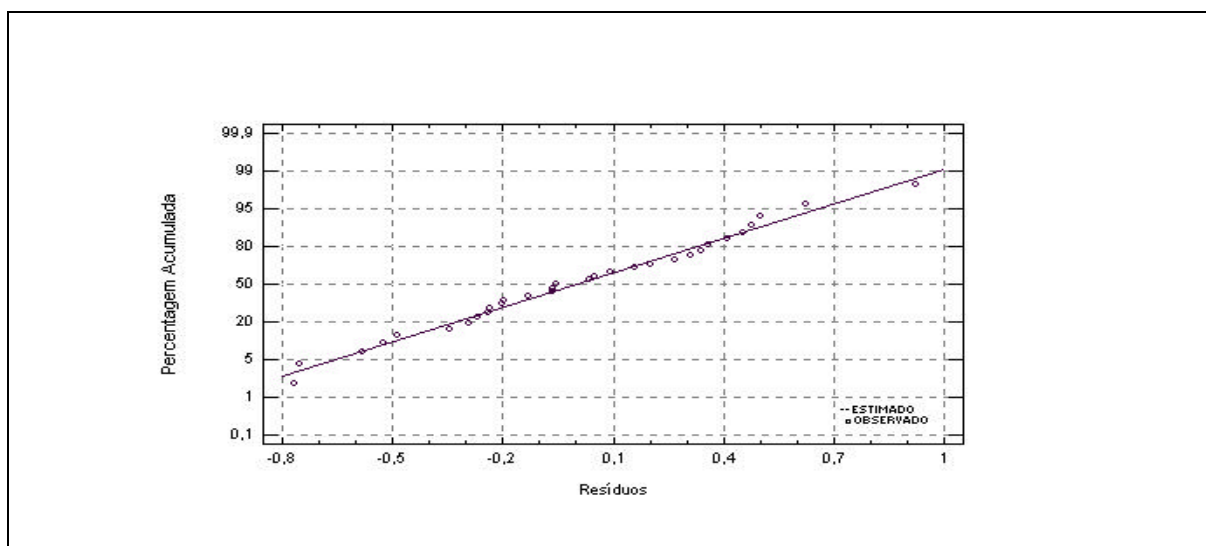


Figura 4.32. Gráfico de probabilidade normal dos resíduos dos custos estimados de emissários.

A probabilidade normal dos resíduos, plotada acima indica, que o modelo não apresenta deficiências, pois os resíduos seguem uma distribuição normal.

Na análise dos resíduos, não há dados espúrios a considerar.

A seguir são apresentados dois gráficos que mostram a relação entre os dados de custo da amostra e os estimados pelo modelo.

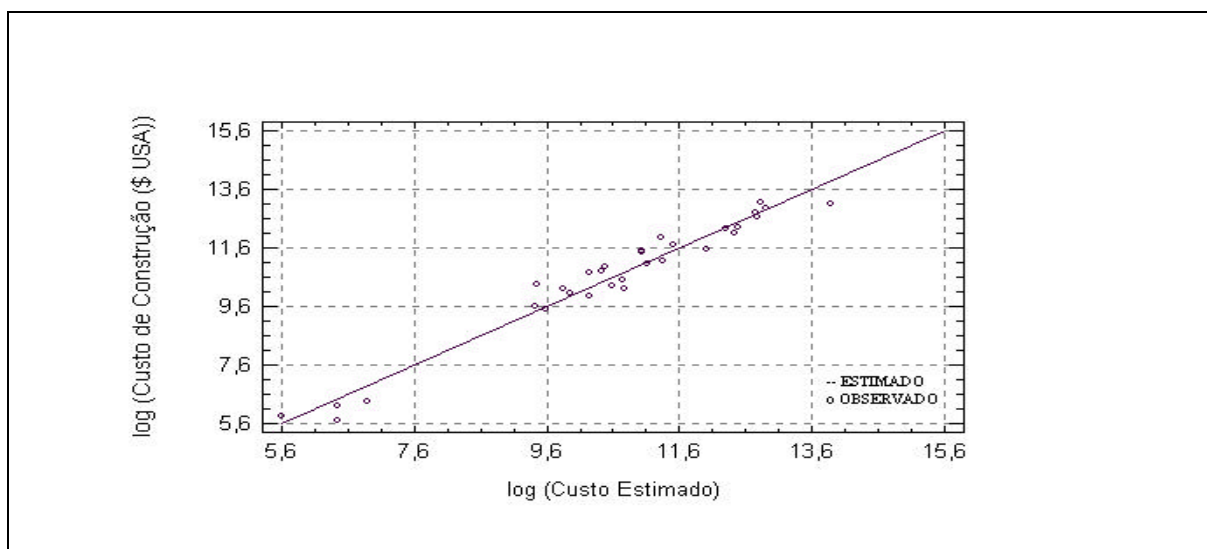


Figura 4.33. Custos estimados em função de custos observados de emissários.

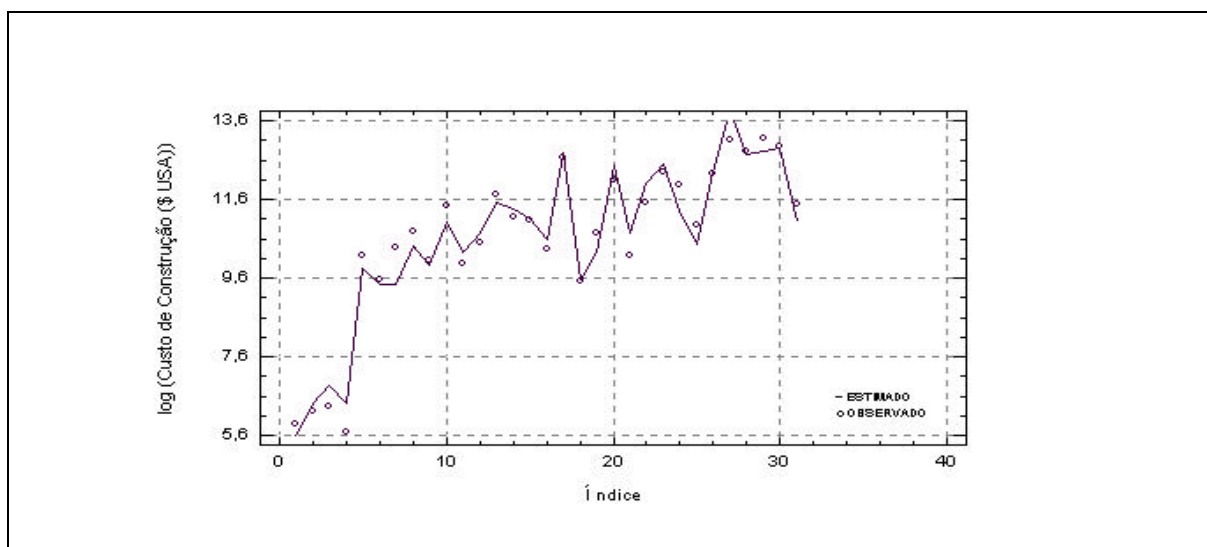


Figura 4.34. Custos estimados e observados de emissários relacionados aos índices (numeração de ordem dos dados).

Os gráficos das Figuras 4.33. e 4.34. confirmam o alto valor obtido para o coeficiente de determinação, isso significa que vazão de projeto e a extensão de emissário podem explicar mais de 95% da variação dos custos de emissários.

Concluída a análise de regressão múltipla de custo de emissário, o modelo tem a seguinte transformada de sua equação:

Transformada da equação:

$$\log(Y_4) = 2,31766 + 1,04354.\log(L_4) + 0,548323.\log(X_4) \quad (11)$$

Para facilitar o uso da equação, sua forma direta:

$$Y_4 = 10,152 L_4^{1,044} X_4^{0,548} \quad (12)$$

4.5 Modelo de Custo de Estação de Tratamento de Esgoto Sanitário

Há vários tipos de estações de tratamento. Dentre os disponíveis na CASAN são coletados dados de apenas 4 projetos de estações de lodos ativados em condições de sistematização e normalização.

O custo da estação é caracterizado, no caso de lodos ativados, pelos seguintes serviços: locação, limpeza do terreno, escavação, aterro, escoramentos, esgotamento, execução de estruturas de concreto, acabamentos, impermeabilização, aquisição, montagem de equipamentos, de tubos, peças e conexões e testes de funcionamento.

Vários dados de projeto de estação de tratamento relacionáveis ao seu custo podem ser selecionados. Entretanto, os mais fáceis e simples de obter já nas primeiras fases são a população geradora dos esgotos e a vazão de projeto. Essas variáveis são consideradas direcionadoras de custo. Observe-se que quando elas aumentam de valor, o custo da estação também aumenta.

O modelo de custo, definido para essa parte de sistema, pode, portanto, ser explicado pelas variáveis independentes: vazão de projeto medida em litros por segundo (l/s) e a população contribuinte de esgoto medida em número de habitantes (hab.).

A amostra dos dados históricos normalizado e consistente para formulação da relação paramétrica de custo da estação é mostrada na tabela a seguir.

Tabela 4.15. Dados de projetos de estação de tratamento.

Número de Ordem	Vazão (l/s)	População (hab.)	Custo (R\$ 07/00)	Custo (\$ USA)
1	27,8	16.000	330.471,80	185.345,90
2	9,2	6.000	751.634,82	421.556,30
3	132,8	55.100	2.993.944,79	1.679.161,00
4	81,0	28.398	3.092.010,84	1.734.162,00

Fonte: Projetos arquivados na Companhia Catarinense de Águas e Saneamento (CASAN)

Observando os dados acima, é possível identificar que existe algum fator determinante na variabilidade dos custos. Os dados da primeira e segunda linhas da tabela, mostram que os valores da vazão e da população da primeira são maiores que da segunda, enquanto que os valores de custos ocorrem ao contrário. Veja-se que ocorre o mesmo com a terceira e quarta linha. A causa dessa variabilidade está na falta de padronização dos projetos, principalmente em relação aos equipamentos especificados.

A equação matemática, que melhor expressa a relação e seu conseqüente o modelo paramétrico, é a seguir apresentada:

$$Y_5 = K_5 L_5^b X_5^a \quad (13)$$

Onde:

Y_5 = custo de estação de tratamento de esgoto sanitário (\$ USA);

K_5 = constante de custo para estação;

L_5 = população contribuinte de esgotos (hab.);

b = expoente de L_5 que representa o fator de economia de escala da população;

X_5 = capacidade da estação de tratamento (vazão de projeto) (l/s);

a = expoente de X_5 que representa o fator de economia de escala da capacidade da estação de tratamento.

A tabela a seguir mostra os resultados estatísticos da regressão múltipla de ajustamento do modelo.

Tabela 4.16. Ajuste do modelo de custo de estações de tratamentos.

Parâmetros	Estimativa	Erro Padrão	t estatístico	Nível de Significância
K_5	502.038,0	107.589,4	4,6662	0,1343
a	53.478,4	18,8	-5,7517	0,1095
b	-108,0	7.151,4	7,4780	0,0846

Os valores dos testes, na tabela acima, mostram que os fatores de economia de escala (expoentes) não são significativamente diferentes de zero. Isto é, ambas as variáveis independentes, capacidade e população, podem não ser relevantes na avaliação de custo de estação de tratamento de esgoto sanitário.

Os testes de significância da hipótese nula das variáveis direcionadoras, analisados pela distribuição de Student, estão representados nas duas últimas colunas da tabela 4.16 com níveis de significância muito altos, isto é, muito diferentes de zero, para o expoente *a* e para o expoente *b*. Isso indica que a hipótese dos fatores de escala dessas variáveis, definidos pela regressão, serem zero não é nula. Se o expoente de uma variável é zero, qualquer valor que ela assuma terá como resultado sempre o número um. Assim, a variável não influenciará os valores de custo resultantes do modelo.

Os resultados do erro padrão, o qual mede a precisão dos fatores de escala, estimados em termos absolutos, mostram uma variação muito grande para o fatores de escala e para a constante. Esse resultado demonstra menor precisão na determinação desses fatores, com conseqüências significativas nos resultados do modelo.

A análise de variância do modelo é mostrada abaixo, na tabela 4.17.

Tabela 4.17. Análise da variância do ajuste do modelo.

Fonte	Somatório dos Quadrados	Graus de Liberdade	Quadrado Médio	Razão F	Valor p
Modelo	$198,540 \times 10^7$	2	$992,697 \times 10^7$	76,25	0,0807
Resíduo	$130,172 \times 10^5$	1	$130,172 \times 10^5$		
Total	$199,840 \times 10^7$	3			

A hipótese testada mostra que a Razão F só é significativamente maior que um dentro do nível de significância superior a 5%. Isso significa que o modelo não tem boa aderência a tendência dos dados observados num intervalo de confiança de 95%.

Outros dados estatísticos para o modelo:

Tamanho da amostra	4
Dados desconsiderados	0
Coeficiente de determinação (R^2)	99,67%
Coeficiente de determinação ajustado (R^2 ajustado)	98,05%
Desvio padrão da estimativa	114.093,3

Embora o coeficiente de determinação, mostrado acima, indique que o modelo explana pelo menos 98% da variabilidade dos custos, isto não é muito significativo, dado a alta variabilidade demonstrada pelos resíduos e principalmente pelo inexpressivo tamanho da amostra. É sabido que quanto menor for o tamanho da amostra menos significativo é o coeficiente de determinação na explicação do modelo.

4.6 Análise dos Resultados dos Modelos

Os resultados dos ajustes das relações paramétricas de custo de cada parte de sistema de esgoto sanitário estão resumidos na Tabela 4.18.

Tabela 4.18. Modelos matemáticos para as relações paramétricas de custo dos componentes de sistema de esgoto sanitário.

Componentes de Sistema	Modelo Matemático	R^2
Rede Coletora	$Y_1 = 669,56 L_1^{0,487} X_1^{1,036}$	96%
Ligações Prediais	$Y_2 = 127,89 L_2^{1,040}$	96%
Estação Elevatória	$Y_3 = 3,29 \times 10^3 L_3^{0,330} X_3^{0,679}$	94%
Emissário	$Y_4 = 10,152 L_4^{1,044} X_4^{0,548}$	95%
Estação de Tratamento	<i>não modelado^(*)</i>	

(*) – Devido a pouca significância nos testes estatísticos das variáveis e de suas relações no modelo

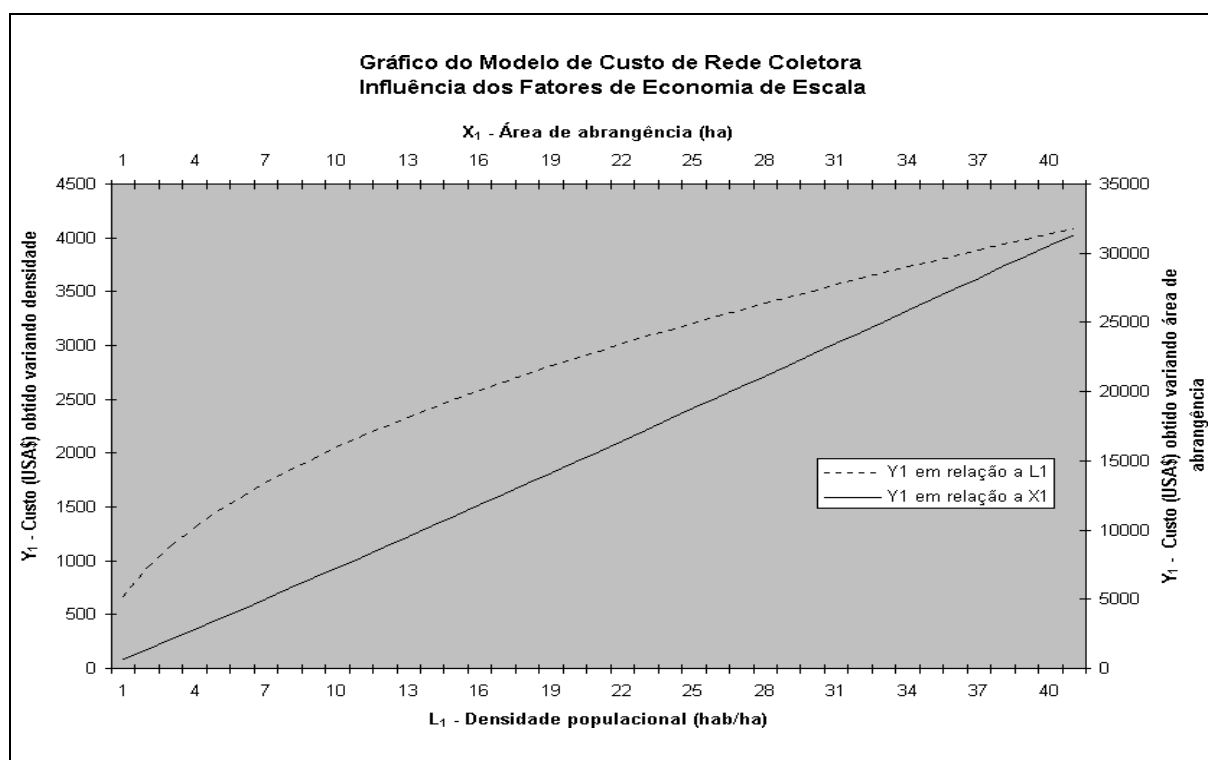


Figura 4.35. Influência dos fatores de economia de escala no modelo de custo de rede coletora.

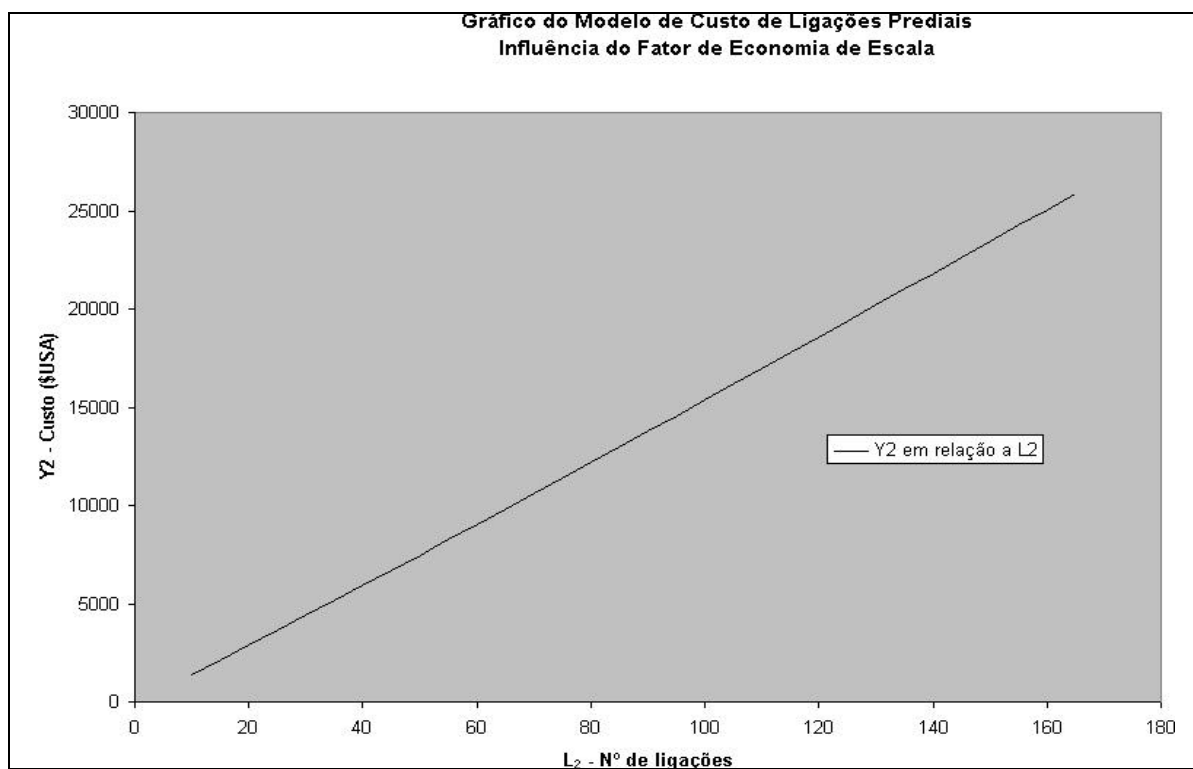


Figura 4.36. Influência do fator de economia de escala no modelo de custo de ligações prediais.

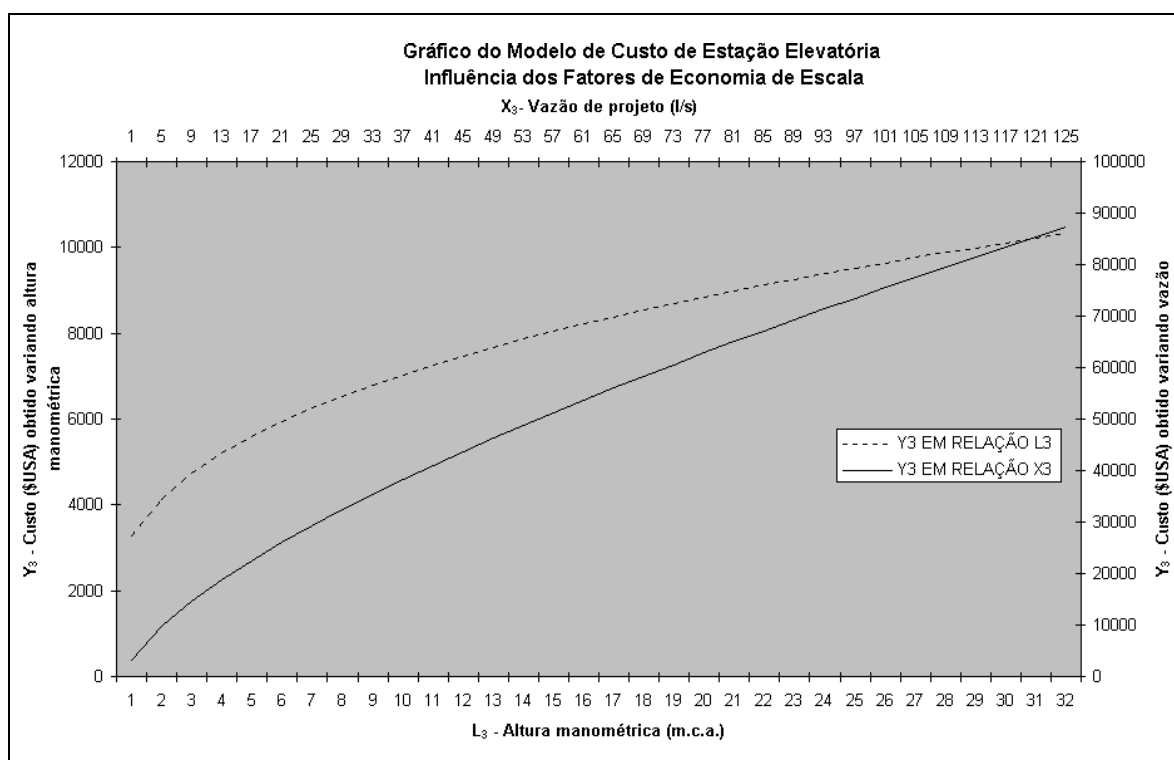


Figura 4.37. Influência dos fatores de economia de escala no modelo de custo de estação elevatória.

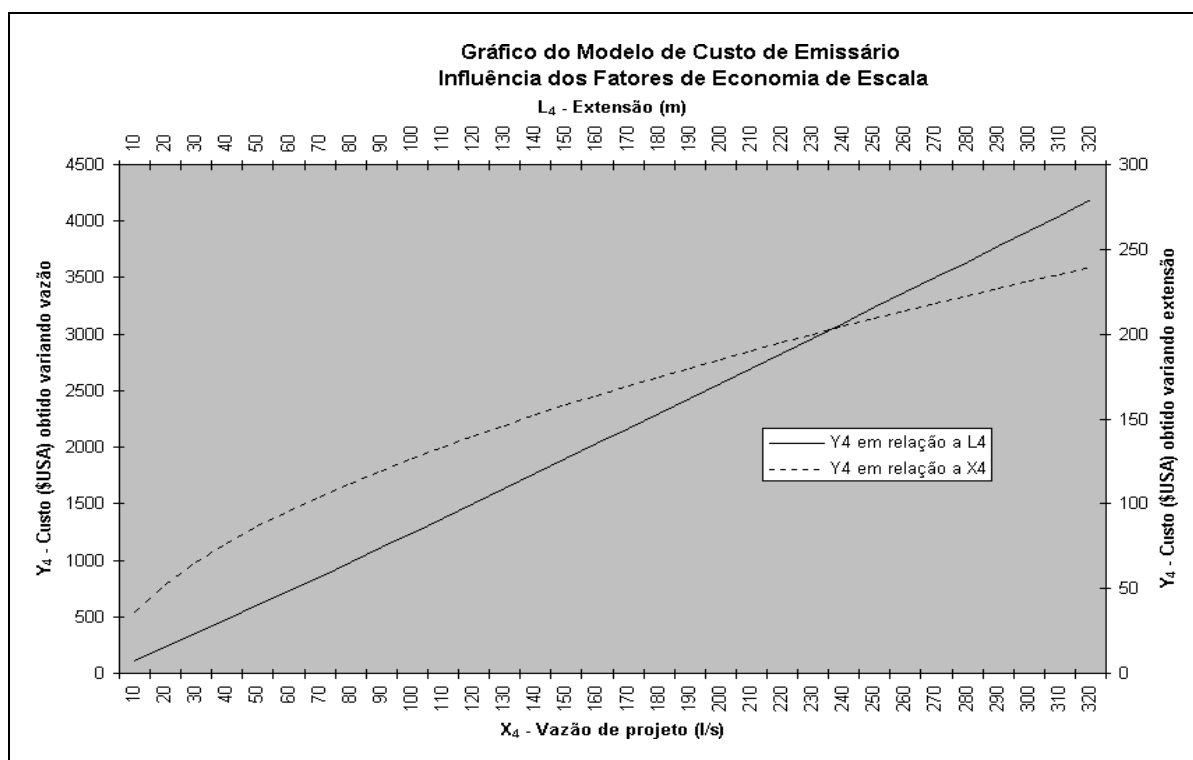


Figura 4.38. Influência dos fatores de economia de escala no modelo de custo de emissário.

Nesses resultados de ajuste dos modelos, constata-se que o fator de economia de escala é significativo para a maioria dos componentes. Os fatores para a capacidade, identificada no modelo da rede coletora pela densidade e no do emissário pela vazão, apresentam relativamente a maior economia de escala. Desse modo, quanto maior a capacidade menores são os incrementos de custos de construção. Já, no caso do modelo da estação elevatória que tem duas variáveis independentes, ocorre que o fator de economia de escala para a capacidade (vazão) não é tão representativo no decréscimo do incremento de custo.

No modelo de custo para rede coletora, conforme figura 4.35, a variável densidade (L_1) mostra mais influência na obtenção de economia de escala que a área de abrangência (X_1), pois o seu efeito é bastante superior à contribuição da área. Isso é mostrado na prática, que para uma mesma área, quanto maior a densidade menor é o incremento de custo. No caso da área de abrangência, o motivo de seu fator ser muito próximo de 1,0, indica que há uma relação diretamente proporcional com seu custo, ou seja, se ela cresce, o custo cresce na razão direta.

Analizando o modelo de ligações prediais, usando o gráfico da figura 4.36, observa-se que o fator de economia de escala da variável explicativa (número de ligações) é bastante próximo de 1,0, indicando que o custo cresce, praticamente, numa razão direta com o número de ligações. Como já foi explicado, o número de ligações é constituído por ligações de 100mm e de 150mm de diâmetro, numa proporção média de 4:1. Portanto, se o número de ligações de 150mm cresce, esse fator tenderá a crescer também. Porém, devido a alta proporção de ligações de 100mm, a influência do número de ligações do outro diâmetro no modelo não é representativa.

O modelo de custo, para estação elevatória, mostra na figura 4.37 que a altura manométrica (L_3) é mais influente no custo que a capacidade (vazão de projeto) (X_3) relativamente a economia de escala, aquela contribui mais que essa para o decaimento do incremento de custo. Assim, os resultados indicam que, em relação ao crescimento da altura manométrica, os custos unitários de construção são fortemente decrescentes, enquanto que em relação ao crescimento da capacidade essa redução não é tão expressiva. Esse modelo, se comparado com o de estação de recalque de água no trabalho de JUNGLES(94), apresenta muita semelhança. Lá,

os fatores são ($a=0,750$ e $b=0,340$), bastante parecidos com os do modelo em análise, onde ($a=0,679$ e $b=0,330$). Embora as tecnologias de construção sejam diferentes, da água para o esgoto, os custos dos equipamentos apresentam semelhança na influência sobre os custos de construção. Portanto, a forma das curvas de custos são similares, alterando-se apenas os níveis de custo (constante K).

No modelo de custo de emissário, foram definidas, também, duas variáveis, extensão (L_4) e vazão de projeto (X_4). No gráfico da figura 4.38, a variável X_4 apresenta uma alta influência na economia de escala dos custos, se comparada com L_4 . Seus fatores $a=0,548$ e $b=1,044$ são de aproximadamente metade um do outro refletidos nos custos, ou seja, o crescimento da vazão influi fortemente no decrescimento do incremento de custo de construção, enquanto a extensão influi numa razão praticamente constante no crescimento do custo.

Relativamente a impossibilidade de modelação para estação de tratamento foram observadas duas causas principais. Uma é o tamanho da amostra, que são apenas quatro conjuntos de dados. A outra é a variabilidade dos custos coletados de estações de tratamento, devido a falta uma padronização nos seus projetos, conforme já observado na avaliação dos resultados.

4.7 Validação dos Modelos

Para avaliar se as relações paramétricas construídas nesse trabalho são boas para predição dos custos de partes componentes de sistema de esgoto sanitário, é necessário avaliar a precisão delas. Assim, são separados aleatoriamente dados independentes do conjunto de dados históricos que geraram os modelos paramétricos. Para cada parte de sistema, são coletados tantos dados quanto disponíveis, desde que o conjunto de dados formadores de cada modelo não sejam considerados insuficientes para sua construção.

A quantidade de pares de dados coletados para validação dos modelos, não é grande suficiente para uma perfeita avaliação deles, haja vista o pequeno número de dados históricos para alguns componentes de sistema de esgoto. Foram coletados para validação dos modelos:

- 04 pares de dados para rede coletora;
- 05 pares de dados para ligações prediais;
- 06 pares de dados para estação elevatória;
- 04 pares de dados para emissários.

A comparação dos resultados de custos, estimados pelos modelos e os coletados dos projetos, pode ser feita observando os valores mostrados nas tabelas 4.23; 4.24; 4.25 e 4.26.

Para efeito de comparação, os dados de custo são saneados dos mesmos efeitos que podem distorcer o padrão estabelecido no processo de seleção dos dados históricos, e normalizados para a mesma data dos modelos, ou seja, para julho de 2000 e convertidos para o dólar americano, a mesma moeda utilizada pelos modelos.

Tabela 4.19. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de rede coletora.

Nome do Sistema	Custo de Projeto (\$ USA)	Custo Estimado (\$ USA)	Variação* (%)	Data do Custo
Joinville (Bacia 3.2)	2.976.040,00	2.859.622,00	3,9	07/2002
Barra da Lagoa	751.834,00	680.239,00	9,5	03/1998
Xanxerê	1.124.860,00	974.617,00	13,4	04/2002
Criciúma (Rio Bonito)	104.287,00	100.589,00	2,6	06/2002

(*)Variação percentual do custo estimado em relação ao de projeto.

Os resultados do modelo acima mostram, embora seja uma amostra reduzida, que ele estimou valores com erro máximo de 13,4% e uma média dos erros de 7,4%

Tabela 4.20. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de ligações prediais.

Nome do Sistema	Custo de Projeto (\$ USA)	Custo Estimado (\$ USA)	Variação* (%)	Data do Custo
Joinville (Bacia 3.2)	843.412,00	801.115,00	5,0	07/2002
Criciúma (A. Sena)	84.555,00	81.894,00	3,1	06/2002
Criciúma (R. Bonito)	44.710,00	41.983,00	6,1	06/2002
Criciúma (C. Rei II)	24.820,00	22.764,00	8,2	06/2002
L. Conceição (Bacia 4)	62.275,00	59652,00	4,2	07/2002

(*)Variação percentual do custo estimado em relação ao de projeto.

O modelo de estimação mostrou um erro máximo de 8,2% e uma média de 5,3% para a diferença em relação ao custo de projeto.

Tabela 4.21. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de estação elevatória.

Nome do Sistema	Custo de Projeto (\$ USA)	Custo Estimado (\$ USA)	Variação* (%)	Data do Custo
Joinville (Bacia 3.2)	200.926,00	247.202,00	-23,1	07/2002
Criciúma (A. S. - ER1)	23.605,00	22.772,00	3,5	06/2002
Criciúma (A.S. – ER2)	9.567,00	8.080,00	15,6	06/2002
Criciúma (C.R. – ER1)	15.653,00	13.661,00	12,7	06/2002
Barra Lagoa (EED)	47.072,00	46.607,00	1,0	03/1998
Bombinhas (ER2/F)	69.510,00	74.448,00	-7,0	07/2001

(*)Variação percentual do custo estimado em relação ao de projeto.

A estimativa do modelo de estação elevatória apresentou um erro máximo de 23,1% e uma média de -0,5% e 10,5% para a média das diferenças absolutas.

Tabela 4.22. Comparativo entre os custos de projetos da CASAN e aqueles estimados pelo modelo de emissário.

Nome do Sistema	Custo de Projeto (\$ USA)	Custo Estimado (\$ USA)	Variação* (%)	Data do Custo
Joinville (Bacia 3.2)	152.167,00	195.223,00	-28,3	07/2002
Criciúma (C.R.–EMER1)	127,00	167,00	-31,9	06/2002
Catanduvás (C.S.K/L.AFAST)	22.882,00	19.755,00	13,6	06/1997
Bombinhas (EM-ER2F)	63407,00	67.117,00	-5,8	07/2001

(*)Variação percentual do custo estimado em relação ao de projeto.

O modelo de emissário estimou valores com erro máximo de 31,9% e média relativa de -13.1% e media absoluta de 19,9%

Os resultados das quatro tabelas anteriores mostram que os modelos paramétricos podem ser usados para estimar custos de construção de partes de sistema de esgoto sanitário através das variáveis direcionadoras desses modelos propostos.

4.7.1 Comparativo para um Sistema entre o Custo de Projeto e o Estimado pelos Modelos

Para avaliar e comparar o custo total de um sistema na validação dos modelos, são utilizados os dados de custo e de seus direcionadores do Sistema de Esgoto Sanitário de Joinville – Bacia 3.2., conforme mostrados na tabela a seguir.

Tabela 4.23. Dados de custo de projeto do S.E.S. de Joinville – bacia 3.2.

Componente	Variável (L)	Variável (X)	Custo de Projeto (R\$ 07/2002)
Rede Coletora	48,01 hab/ha	517,4 ha	6.379.399,20
Ligações Prediais	4456 ligações		1.807.926,17
Estação Elevatória	225,05 l/s	6,3 m.c.a	430.542,18
Emissário	225,05 l/s	740,0 m	326.061,48
Total			8.943.929,03

Fonte: projeto arquivado na Companhia Catarinense de Águas e Saneamento (CASAN)

Com as equações dos modelos de cada componente e as variáveis da tabela 4.23. são obtidos os custos estimados em dólar americano (\$ 07/2000), mostrados

abaixo, juntamente com os custos de projeto, também em dólar.

Tabela 4.24. Custos de projeto e estimados do S.E.S. de Joinville – bacia 3.2.

Componente	Custo de Projeto (\$ 07/2000)	Custo Estimado (\$ 07/2000)	Variação* (%)
Rede Coletora	2.976.040,00	2.859.622,00	3,9
Ligações Prediais	843.411,99	801.115,26	5,0
Estação Elevatória	200.926,00	247.202,00	-23,1
Emissário	152.167,00	195.223,00	-28,3
Total	4.172.544,99	4.103.162,26	1,7

(*)Variação percentual do custo estimado em relação ao de projeto.

Na tabela acima, observa-se que os custos de projeto e os custos estimados apresentam variações entre si que são aceitáveis para os modelos paramétricos. Quando somados, seus valores totais tornam-se bastante próximos, com variação de apenas 1,7%. Isso se deve ao fato de que os custos da rede coletora e das ligações prediais são freqüentemente os itens de maior vulto, como no caso em tela onde representam cerca de 90% do custo total do sistema. São eles, por conseguinte, que melhoram a precisão dos custos totais estimados.

Embora não se disponha de um número maior de projetos disponíveis para comparação, pode-se considerar que os valores estimados dos modelos, quando utilizados para avaliar os custos totais de um sistema, tendem a apresentar um custo estimado total mais próximo do real de projeto.

Para completar o entendimento do leitor, a tabela abaixo mostra os custos de projeto e estimado do sistema de Joinville, em Reais de 07/2002, utilizando os mesmos índices de normalização.

Tabela 4.25. Custo de projeto do S.E.S. de Joinville – bacia 3.2. e custo estimado atualizado

Custo de Projeto (R\$ 07/2002)	Custo Estimado (R\$ 07/2002)
8.943.929,03	8.795.481,35

5 CONCLUSÕES E RECOMENDAÇÕES

5.1 Conclusões do Trabalho

Algumas das conclusões a que se chega estão coerentes com conhecimentos anteriores presentes em outras pesquisas e na visão tradicional existente no setor do saneamento. Outras vêm trazer fatos novos que devem facilitar o trabalho de estimativa de custos neste setor e, em decorrência disso, o desenvolvimento de diversas atividades ligadas a tais estimativas. Dentre elas, o principal destaque é a obtenção de modelos que facilitam as estimativas de custo de:

- Rede Coletora: $Y_1 = 669,56 L_1^{0,487} X_1^{1,036}$, onde Y_1 é o custo (\$USA), L_1 é a densidade populacional (hab/ha) e X_1 é a área de abrangência da rede (ha).
- Ligações Prediais: $Y_2 = 127,89 L_2^{1,040}$, onde Y_2 é o custo (\$USA) e L_2 é o número de ligações.
- Estação Elevatória: $Y_3 = 3,29 \times 10^3 L_3^{0,330} X_3^{0,679}$, onde Y_3 é o custo (\$USA), L_3 é a altura manométrica (m.c.a.) e X_3 é a vazão de projeto (l/s).
- Emissário: $Y_4 = 10,152 L_4^{1,044} X_4^{0,548}$, onde Y_4 é o custo (\$USA), L_4 é a extensão do emissário (m) e X_4 é a vazão de projeto (l/s).

5.1.1 Conclusões Gerais

Os modelos paramétricos de custo desenvolvidos para partes componentes de sistema de esgoto sanitário não necessitam de grande detalhamento de projeto para estimativa de seus custos, pois as variáveis direcionadoras são encontradas já nas primeiras fases de seus estudos. Conseqüentemente, não há necessidade de grandes investimentos em projetos e nem de aguardar o detalhamento deles para obter seus custos de implantação, como se faz para outros métodos tradicionais de estimativa.

Embora o pequeno número disponível de dados para validação, eles fornecem resultados relativamente confiáveis para cada uma das partes de sistema. Isso se considerar, é claro, que as médias das diferenças dos custos reais de projeto e dos estimados não superam a 8% dos custos reais, para rede coletora, 4% para ligações prediais, 14% para estação elevatória e 20% para emissário.

Cabe salientar também que os custos estimados de cada parte de sistema, quando somados configurando o custo total de um sistema, demonstram-no bastante próximo do custo real do projeto.

5.1.2 Conclusões Específicas

O nível de precisão alcançado pelas relações paramétricas de custo de cada parte de sistema de esgoto pode ser considerado bom, se considerar a sua utilização antes do desenvolvimento do projeto, para planejadores, para agentes de financiamentos de saneamento e para projetistas na fase de estudos econômicos.

As variáveis direcionadoras escolhidas resultaram para os modelos coeficientes de determinação superiores a 94%, indicando que os modelos explicam adequadamente os custos de cada componente do sistema.

Embora seja reduzido número disponível de dados para teste dos modelos, os valores estimados para cada componente apresentaram bom nível de precisão, e quando testados configurando um sistema completo, o valor total estimado apresenta-se ainda mais próximo do valor real de projeto.

Na investigação das variáveis explanadoras dos custos, as que melhor os explicam, visando o não detalhamento do projeto, são:

- a área de abrangência, medida em hectare, e a densidade de saturação populacional da área, medida em habitantes por hectare, para o modelo de rede coletora;
- o número de ligações para o modelo de ligações prediais;
- a vazão de projeto, medida em litros por segundo, e a altura manométrica, medida em metros de coluna d'água, para o modelo de estação elevatória;
- e vazão de projeto, medida em litros por segundo, e a extensão de tubulação, medida em metros, para o modelo de emissário.

De acordo com os resultados apresentados da análise estatística, a relação entre a variável dependente e as variáveis independentes resultaram em funções não lineares para todos os modelos. A busca de relações que resultassem em modelos lineares de primeira ordem, demonstraram-se pobres, quando comparado o

coeficiente de determinação, o quadrado do erro médio e o coeficiente C_p de Mallows.

5.2 Como Utilizar os Modelos

Para estimar custos de componentes de sistema de esgoto sanitário é necessário seguir os procedimentos abaixo:

- **Estimativa de custo de rede coletora** – O uso do seu modelo requer a determinação de duas variáveis, uma é a área urbana (X_1) que se quer atender, em hectare, e a outra a densidade populacional de saturação (L_1) dessa área, em habitantes por hectare. A primeira variável é obtida através de mapas ou levantamentos topográficos da área e a segunda, através de plano diretor ou de estudo populacional da área urbana. De posse dos valores das variáveis, a equação do modelo pode ser utilizada para obter-se o custo estimado da rede coletora. O seu valor calculado em dólares americano deve ser transformado em real, utilizando-se as tabelas de câmbio comercial existentes nas páginas de economia de qualquer jornal de grande circulação ou revistas especializadas em economia.
- **Estimativa de custo de ligações prediais** – O modelo de custo é função de apenas uma variável, que é o número de ligações (L_2). A obtenção do número de ligações prediais à rede coletora de esgoto sanitário de uma determinada área urbana pode ser feita de duas maneiras. Uma através do cadastro de ligações de abastecimento de água daquela área e a outra usando o levantamento cadastral do município onde consta a área. Definido o número de ligações prediais, a equação do modelo permite estimar o custo das ligações prediais da área que se quer atender com rede de esgoto sanitário. Do mesmo modo que a estimativa da rede coletora, o custo é calculado em dólares e deve ser transformado em real.
- **Estimativa de custo de estação elevatória** – O cálculo do custo é feito usando o modelo de custo de estação elevatória, cuja a equação é função de duas variáveis. Uma, é a vazão de projeto (X_3) da estação, em litros por segundo. Essa pode ser obtida através do estudo de concepção do esgotamento sanitário da área urbana ou bacia que se quer esgotar e

transferir para outra área, bacia de contribuição ou destino final. A outra variável é a altura manométrica (L_3), em metros de coluna d'água, obtida pelo cálculo da diferença de cota geométrica, entre o nível topográfico da estação elevatória e o nível do local de destino do esgoto bombeado. De posse dos valores das variáveis, entra-se na equação do modelo para se obter a estimativa de custo da estação elevatória objeto de avaliação.

- **Estimativa de custo de emissário** - É feita a partir do modelo de custo de emissário, cuja equação é função da vazão de projeto (X_4), transportada pela tubulação, em litros por segundo, e da extensão (L_4) dessa tubulação, em metros lineares. Ambas são obtidas no estudo de concepção de esgotamento sanitário da área urbana ou bacia que se quer esgotar. A vazão pode ser a mesma da estação elevatória ou então a vazão no tubo, por gravidade, que deva ser transportada de um ponto para outro. A extensão, por seu lado, pode ser obtida medindo-se em mapas ou levantamentos topográficos. Com os valores de X_4 e L_4 determinados, a estimativa de custo de emissário pode ser calculada e transformada em real na data que se queira a avaliação.

5.3 Recomendações para Pesquisas Futuras

A seguir apresentam-se proposições para pesquisas futuras que podem acrescentar informações ao estudo da utilização de relações paramétricas de custo no setor do saneamento, de acordo com as conclusões alcançadas dentro desse trabalho. São elas:

- análise dos orçamentos de projetos geradores dos dados de custo, visando buscar maior homogeneidade dos custos, diminuindo assim a variabilidade encontrada nos modelos e o conseqüente erro padrão da estimativa;
- avaliação de modelos paramétricos levando em conta as diferenças regionais de custo, de materiais, de topografia e de geologia;
- busca de modelos existentes para os vários tipos de estações de tratamento de esgoto que possam ser calibrados para usá-los na estimativa de seus custos;

- análise de outras variáveis que possam explicar melhor o incremento dos custos relativos a “verticalização” da área de atendimento do projeto no modelo de ligações prediais;
- estudo de índice que possa melhor representar a correção da inflação, em substituição aos existentes não específicos, visando a normalização dos dados de custo de saneamento coletados para os modelos;
- reavaliação dos modelos paramétricos para rede coletora e para ligações prediais baseados em amostras de maior tamanho, visando assim melhorar a qualidade de ajuste deles.

6 REFERÊNCIAS

ABRANOVIC, W. A. **Statistical thinking & data analysis methods for managers**. Reading, MA: Addison-Wesley Publishing Company, 1997.

AIR FORCE MATERIAL COMMAND (AFMC). **Volume II Aeronautical**. Cost estimating handbook series.. Fairborn: The Analytic Sciences Company, cerca de 1992.

ANTUNES JÚNIOR, J.A.V., **Em direção a uma teoria geral do processo na administração da produção: uma discussão sobre a possibilidade de unificação da teoria das restrições e da teoria que sustenta a construção dos sistemas de produção com estoque zero**. Doutorado em Administração. Escola de Administração/Universidade Federal do Rio Grande do Sul. Porto Alegre/RS, 1998.

ASHER, H. **Cost-quantity relationships in the airframe industry**, R-291. Santa Monica: Rand Corporation, 1956.

BENJES, Henry H. et al. **Cost estimating manual – combined sewer overflow storage and treatment**. El Dorado Hills: U.S. EPA, 1976.

BIRKLER, J. L., GARFINKLE, J. B. **Regression diagnostics in practice-experiences from modeling jet engine costs**, P-6896.. Santa Monica: Rand Corporation, 1983.

BIRKLER, J., LARGE, J., SMITH, G. et al. **Reconstituting a production capability**, MR-273-ACQ. Santa Monica: Rand Corporation, 1993.

BOOTHROYD, Geoffrey. **Cost estimating at the product design stage**, paper. In: 1998 JOINT ISPA/SCEA INTERNATIONAL CONFERENCE. Toronto: ISPA, 1998.

BORNIA, A.C., **Mensuração das perdas dos processos produtivos: uma abordagem metodológica de controle interno**. Doutorado em Engenharia de Produção. Universidade Federal do Santa Catarina. Florianópolis/SC, 1995.

BORNIA, A.C.; SANTOS, N.J.dos, **Custeio baseado em atividades – ABC: aspectos da análise dos processos e atividades**. In: Congresso Brasileiro de Gestão Estratégica de Custos, 4. 1997. Anais, p. 51-57. Petrônio Cunha dos Santos/MG, 1997.

COKINS, Gary. **Why is traditional accounting failing managers?** paper. In: JOINT ISPA/SCEA INTERNATIONAL CONFERENCE. Toronto: ISPA, 1998.

COLLINS, Jim. **Parametric cost estimating initiative**, presentation. Price Symposium. Albuquerque, 1996.

COMPANHIA DE SANEAMENTO BÁSICO DO ESTADO DE SÃO PAULO – SABESP. **Estudo de curvas de custos de empreendimentos**. São Paulo, 1998.

COYOTE Sky. Software Engineer, Intergalactic Reality. **Arbitrary n-parameter linear and non-linear least-squares regression analysis of 1 and 2 dimensional data or curve-fitting**. [on line] 1994/1995. [citado 06 jun. 2001] Disponível na World Wide Web: <<http://www.intergalactic.com/seam95/seam95.html>>

DEPARTMENT OF DEFENSE (DOD). **Parametric estimating handbook**. 2nd edition. Washington, 1999.

EARLE, George A., FARREL JR., Paul. **A mathematical model for estimating sewer costs**. In: NEW ENGLAND WATER ENVIRONMENT ASSOCIATION ANNUAL CONFERENCE. Boston: Environment One Corporation, 1997.

FERENS, Daniel V. **Software size: quo vadis?**. National estimator. Alexandria, p. 43 - 54, winter 1999.

FOUSSIER, Pierre. **Is regression analysis the right tool for preparing CER's ?**, paper. In: JOINT ISPA/SCEA INTERNATIONAL CONFERENCE. Toronto: ISPA, 1998.

FULLER, James L., ARMSTRONG, Nathalene. **How to analyze and estimate cost**. In: JOINT ISPA/SCEA INTERNATIONAL CONFERENCE. San Antonio: ISPA, 1999.

GILLOT, S. et al. **Optimization of wastewater treatment plant design and operation using simulation and cost analysis**, paper. Belgium, University of Gent, 1999.

GUMERMAN, Robert C., CULP, Russell L., HANSEN, Sigurd P. **Estimating water treatment costs**: volume 1, summary. Santa Ana: U.S. EPA, 1979.

GUMERMAN, Robert C., CULP, Russell L., HANSEN, Sigurd P. **Estimating water treatment costs**: volume 2, cost curves applicable to 1 to 200 mgd treatment plants. Santa Ana: U.S. EPA, 1979.

HARNETT, D. L., SONI, A. K. **Statistical methods for bussines & economics**. 4th edition. Reading: Addison-Wesley Company, 1991.

JUNGLES, Antônio Edésio. **Análise de alternativas de expansão de capacidade dos sistemas urbanos de abastecimento de água em Santa Catarina**. Florianópolis, 1994. 217p. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, 1994.

KING, E. W. **An auditable parametric based estimating system**, presentation. In: JCMS CONFERENCE. Baltimore, 1998.

KLIEMANN NETO, F. J., **Custos Industriais** – Apostila de Custos Industriais. Universidade Federal de Santa Catarina. Florianópolis, SC. 1986.

KLIEMANN NETO, F.J.; ANTUNES JÚNIOR, J.A.V., **Proposta de um processo de custeio para sistemas “Just in Time” de produção**. PPGA/Universidade Federal do Rio Grande do Sul. Porto Alegre, RS, 1990.

KLIEMANN NETO, F.J.; MULLER C.J., **A mudança dos sistemas de custeio em ambientes modernos de manufatura: um estudo de caso**. In: I Congresso Brasileiro de Gestão Estratégica de Custos. São Leopoldo, RS, 1994.

LARGE, J. P. **Development of parametric cost models for weapon systems**, P-6604. Santa Monica: Rand Corporation, 1981.

LINECK, Thomas S., GUMERMAN, Robert C., CULP, Russel L. **Estimating water treatment costs**: volume 4, computer user's manual for retrieving and updating cost data. Santa Ana: U.S. EPA, 1979.

LONG, John A. **Parametric cost estimating in the new millennium**, paper. In: SCEA -2000 ANNUAL CONFERENCE PAPERS. Manhattan Beach Marriott: SCEA, 2000.

LOSSO, Iseu Reichmann. **Utilização das características geométricas da edificação na elaboração de estimativas preliminares de custos: estudo de caso em uma empresa de construção**. Florianópolis, 1995, 146p. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, 1994.

MADRID, George A. **An unmanned spacecraft subsystem cost model for advanced mission planning**, paper. In: 1998 JOINT ISPA/SCEA INTERNATIONAL CONFERENCE. Toronto: ISPA, 1998.

MANDARANO, Mazzini. **Estimativa de custo para implantação dos serviços de água e esgotos**. In: VI CONGRESSO BRASILEIRO DE ENGENHARIA SANITÁRIA. São Paulo, ABES, 1970.

MARTINS, E., **Contabilidade de custos**. 7ª ed. Editora Atlas. São Paulo/SP, 2000.

MENÉNDEZ, Jose L. S., MORENO, Fernando J. **Análisis de Valor en la F-100. Ingeniería de sistemas aplicada**, monografias, cap. 7, pag. 91-102.. Madrid: Publicaciones de Ingeniería de Sistemas, Isdefe, Ministerio de Defensa, 1995.

MOTULSKY, Harvey J. **Analyzing data with graphpad prism**. [on line] 1999. [citado em 05 jun. 2001] Disponível na World Wide Web: <<http://www.graphpad.com>.

MOTULSKY, Harvey J. **The prism guide to interpreting statistical results**. [on line] 1999. [citado em 05 jun. 2001] Disponível na World Wide Web: <<http://www.graphpad.com>.

OTERO, Juliano Araujo. **Análise paramétrica de dados orçamentários para estimativa de custos na construção de edifícios: estudo de caso voltado para a questão da variabilidade**. Florianópolis, 2000, 214p. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, 2000.

PANZETER, Andrea Angela. **A methodology for modeling the cost and duration of concrete highway bridges**. Purdue University , 1993. 161p. Thesis (Doctor of Philosophy in Civil Engineering) – Purdue University, 1993.

POLLOCK, C. D. **Parametric primer**. Maxwell AFB: AUL/LDEX, 1981.

REILLY, James, GOTTLIEB, Paul. **Estimating costs for wastewater collection and treatment under various growth scenarios**, paper. In: THIRD INTERNATIONAL CONFERENCE ON COMPUTERS IN URBAN PLANNING AND URBAN MANAGEMENT. Atlanta, New Jersey Office of State Planning, 1993.

REILLY, James, GOTTLIEB, Paul. **Projection costs for roads under various growth scenarios**, paper. In: THIRD INTERNATIONAL CONFERENCE ON

COMPUTERS IN URBAN PLANNING AND URBAN MANAGEMENT. Atlanta, New Jersey Office of State Planning, 1993.

RICHARD P. ARBER ASSOCIATES. **What's it going to cost?** Arber News, Denver, Colorado, v.VII, artigo 2, outono, 1998.

RICHTER, Carlos A. **Estimativa de custo de obras de abastecimento de água.** Curitiba, 1978.

ROBBINS, J. L., DANEMAN, J. C. **Parametric estimating & the stepwise statistical technique.** National estimator. Alexandria, p. 24 –34, ago. 1999.

ROBBINS, J. L., SMITH, V. F. **Parametric cost estimating – enjoying new popularity & making new acquaintances,** paper. In: 1999 JOINT ISPA/SCEA INTERNATIONAL CONFERENCE. San Antonio, ISPA, 1999.

SHEKIN, D. J. **Handbook of parametric & non-parametric statistical procedures.** Boston: CRC Press LLC., 1997.

SOCIETY OF COST ESTIMATING AND ANALYSIS (SCEA). **SCEA glossary and definitions.** Washington: SCEA, 1991.

STEVENSON, William J. **Estatística aplicada à administração;** tradução Alfredo Alves de Farias. São Paulo: Harper & Row do Brasil, 1981

STOCKBURGER, David W. **Introductory statistics: concepts, models, and applications.** Southwest Missouri State University. [on line 1996] [citado em 02 jul. 2001] Disponível na World Wide Web: <<http://www.psychstat.smsu.edu/introBook/sbk16m.htm>.

TEOLOGLOU, George – Toussaint. **Classes of cost estimates.** In: ISPA/SCEA INTERNATIONAL CONFERENCE. San Antonio, ISPA, 1999.

TEXAS WATER DEVELOPMENT BOARD. **Cost estimating procedures:** TWDB region H. [citado em 23 nov.2001] Disponível na World Wide Web: http://www.twdb.state.tx.us/rwp/h/Submitted_Files/TWDB%20Reports/Task%205/Appendix%205B.doc

TOURASSE, Enio, COUTINHO, Ataulpho. **Custo de obras de abastecimento de água e esgotos sanitários.** In: III CONGRESSO BRASILEIRO DE ENGENHARIA SANITÁRIA. Curitiba, ABES, 1965.

U.S. ENVIRONMENTAL PROTECTION AGENCY. **Construction costs for municipal wastewater conveyance systems**: 1973 – 1979. Washington: U.S. EPA, 1981.

U.S. ENVIRONMENTAL PROTECTION AGENCY. **Construction costs for municipal wastewater treatment plants**: 1973 – 1978. Washington: U.S. EPA, 1980.

VOLDASE, Iva S. **Parametric cost models: the strategist manager 2000 tool**. In: SCEA NATIONAL CONFERENCE. Manhattan Beach, SCEA, 2000.

WALTER, T. James. **I think, therefore I estimate**. In: ISPA/SCEA INTERNATIONAL CONFERENCE. San Antonio, ISPA, 1999.